

# STARS

University of Central Florida  
**STARS**

---


Electronic Theses and Dissertations, 2004-2019

---

2006

## Class Size Reduction: Is It Worth The Cost? a Meta-analysis Of The Research

Christopher Camacho  
*University of Central Florida*

 Part of the [Curriculum and Instruction Commons](#)  
Find similar works at: <https://stars.library.ucf.edu/etd>  
University of Central Florida Libraries <http://library.ucf.edu>

This Doctoral Dissertation (Open Access) is brought to you for free and open access by STARS. It has been accepted for inclusion in Electronic Theses and Dissertations, 2004-2019 by an authorized administrator of STARS. For more information, please contact [STARS@ucf.edu](mailto:STARS@ucf.edu).

---

### STARS Citation

Camacho, Christopher, "Class Size Reduction: Is It Worth The Cost? a Meta-analysis Of The Research" (2006). *Electronic Theses and Dissertations, 2004-2019*. 829.  
<https://stars.library.ucf.edu/etd/829>



CLASS SIZE REDUCTION: IS IT WORTH THE COST?  
A META-ANALYSIS OF THE RESEARCH

by

CHRISTOPHER M. CAMACHO  
B.S. University of Central Florida, 1997  
M.S.Ed. Nova Southeastern University, 2000

A dissertation submitted in partial fulfillment of the requirements  
for the degree of Doctor of Education  
in the Department of Educational Studies  
in the College of Education  
at the University of Central Florida  
Orlando, Florida

Spring Term

2006

Major Professor: Larry C. Holt

© 2006 Christopher M. Camacho

## ABSTRACT

The purpose of this study was to investigate whether ethnicity, gender, grade level and content area mediate the relationship between class size and student achievement. Twenty six educational research studies were collected for this meta-analysis. A meta-analytical approach using like data sets were used to report the most accurate information. Fixed and random effect models were used to ensure the distribution across different studies. A total of three studies were meta-analyzed for this research.

The studies included in this research examined class size and student achievement for students in grades K-7. This research examined whether there was a mediating effect on ethnicity, gender, grade level, and content area in the class size and student achievement studies collected and correlated. The results indicate that smaller class size does have a positive impact on student achievement when mediated by ethnicity, grade level, and content area. When examining ethnicity as a mediating factor, a stronger correlation exists for minority students than for whites. When examining grade level as a mediating factor results indicate that a stronger correlation exist for lower grades K-3 than upper grades 4-7. A stronger correlation is present in the content area of reading as compared to other content areas when it was used as a mediating factor. When examining gender as a mediating factor, the largest effect sizes were reported for females in reading as opposed to males in math, both for first grade. These results correlate with those presented in other mediator sections of the study. The results found in this research will contribute to the literature concerning class size and student achievement and will help inform educational policy makers regarding class size as an educational issue.

To Angela and CJ

I would have never made it without your love and support.

## ACKNOWLEDGMENTS

This project was completed under the guidance of my dissertation committee who spent countless hours providing support and insight. First, I wish to thank Dr. Larry Holt, committee Chair and advisor for the constant encouragement and friendship since the very beginning. In addition, I thank Dr. David Boote, and Dr. Steven Sivo for their support and interest in this research topic. Finally, I wish to thank Dr. John Edwards for supporting me through my professional career as an educator.

## TABLE OF CONTENTS

LIST OF FIGURES .....	viii
LIST OF TABLES .....	ix
CHAPTER ONE: INTRODUCTION.....	1
Background and Significance .....	1
Statement of the Purpose .....	4
Question of the Study.....	5
Definition of Terms.....	6
Assumptions and Limitations .....	7
Methodology.....	9
Research Design.....	9
Population .....	9
Instruments.....	9
Data Collection .....	9
CHAPTER TWO: REVIEW OF LITERATURE.....	12
Class size reduction and Educational Policy .....	12
Defining Class Size.....	14
Past Class Size Reduction Study Results.....	15
Conflicting Arguments with Class Size Reduction .....	16
Universities and Class Size Reduction .....	17
Standards-Based Accountability and Class Size Reduction .....	18
Direct and Indirect Costs Associated with Class Size Reduction.....	19
Effects on Classroom Instruction.....	20
Effects on Student and Teacher Interaction.....	22
Class Size Reduction and Student Achievement .....	25
State Class Reduction Programs .....	27
Tennessee’s Project STAR .....	27
California’s Class Size Reduction (CSR).....	29
Wisconsin’s SAGE .....	30
Arizona’s Class Size Reduction.....	33
Differences in State Class Size Reduction Programs .....	33
Financial Components Associated with Class Size Reduction.....	34
Summary .....	41
CHAPTER THREE: METHODOLOGY .....	42
Introduction.....	42
Research Design.....	42
Subjects.....	43
Procedures.....	44
Instrument .....	44
Data Collection .....	45
Mediators and Statistical Requirements.....	45
Data Analysis.....	46
CHAPTER FOUR: RESEARCH OF FINDINGS .....	47

Overview.....	47
Stem-and-Leaf Plot Representation of Compressed Effect Sizes.....	52
Funnel plot.....	54
Duval and Tweedie's Trim and Fill.....	55
Overall effect of class size on achievement.....	59
Mediator Results: Ethnicity.....	61
Mediator Results: Gender.....	63
Mediator Results: Grade level.....	66
Mediator Results: K-3 vs. Grades 4-6.....	68
Mediator Results: Grades K-3 vs. Grade 7.....	70
Mediator Results: Grades 4-6 vs. Grade 7.....	72
Mediator Results: Content Area.....	74
Mediator Results: Math vs. Reading.....	77
Mediator Results: Math vs. Science.....	79
Mediator Results: Math vs. Social Studies.....	81
Mediator Results: Reading vs. Science.....	83
Mediator Results: Reading vs. Social Studies.....	85
Mediator Results: Science vs. Social Studies.....	87
<b>CHAPTER FIVE: SUMMARY, DISCUSSION, RECOMMENDATIONS, AND</b>	
<b>CONCLUSION.....</b>	<b>88</b>
Summary of Findings.....	88
Discussion.....	90
Limitations to Generalizability.....	91
Publication Bias.....	93
Recommendations for Future Research.....	94
Conclusion.....	95
<b>REFERENCES.....</b>	<b>97</b>



LIST OF FIGURES

Figure 1 Funnel Plot..... 54

## LIST OF TABLES

Table 1 Studies included in the Class Size Reduction-Student Achievement Meta-Analysis .....	51
Table 2 Collapsed Effect Sizes Obtained from the Studies .....	52
Table 3: Imputed Means to the left - Fixed effects model .....	57
Table 4 Overall Effect of Class Size on Student Achievement .....	58
Table 5 Mediator Results: Ethnicity .....	60
Table 6 Mediator Results: Gender .....	62
Table 7 Mediator Results: Grade level .....	65
Table 8 Mediator Results: Grades K-3 vs. Grades 4-6 .....	67
Table 9 Mediator Results: K-3 vs. Grade 7 .....	69
Table 10 Mediator Results: Grades 4-6 vs. Grade 7 .....	71
Table 11 Mediator Results: Content Area .....	73
Table 12 Mediator Results: Math vs. Reading.....	76
Table 13 Mediator Results: Math vs. Science .....	78
Table 14 Mediator Results: Math vs. Social Studies .....	80
Table 15 Mediator Results: Reading vs. Science.....	82
Table 16 Mediator Results: Reading vs. Social Studies .....	84
Table 17 Mediator Results: Science vs. Social Studies .....	86

## CHAPTER ONE: INTRODUCTION

### *Background and Significance*

Improving student achievement, regardless of an individual's educational views, has become a driving force in the education field. No longer can schools operate without having accountability measures placed on its teachers, administrators, and students. Along with school accountability, several different philosophies have evolved to determine how to improve teacher accountability and more importantly, student achievement.

Legislation has played a critical role in establishing educational standards and goals. In recent years, legislation has impacted education in many different forms. For example, a legislation-driven accountability movement aimed at improving student achievement has led to a great deal of high-stakes. The recent legislation to emerge in several states is a class size amendment, also known as class size reduction, that sets limits on class size at the various levels of schooling. According to the Florida Department of Education (2005), this legislation is entitled Article IX, Section 1 of the Florida Constitution. This Florida legislative measure has taken shape in the efforts to improve student achievement. This legislative goal involves establishing student class size limits, at the various grade levels, to ensure student success. More specifically, the amendment states the maximum number of students who are assigned to each teacher who is teaching in public school classrooms for pre-kindergarten through grade 3 cannot exceed 18 students. In addition, the maximum number of students, under the same conditions, for grades 4 through 8 cannot exceed 22. Furthermore, for grades 9 through 12, under the same conditions, the number of students cannot exceed 25.

Article IX, Section of the Florida Constitution has been termed class size reduction. Many different legislative efforts aimed at improving student achievement have raised several questions over the years. These efforts, especially class size reduction, have led to a very important question, “How will this actually impact student achievement? If it does, where does it impact most?” The purpose of this study is to perform a meta-analysis on available research pertaining to student achievement as related to class size reduction.

Smaller class size, whether part of the private or public school sector, has been a proposal thought to improve student achievement. According to Flake, Von Dohlen, and Gifford (1995), class sizes have decreased over the past few decades. In their research, they note a drop in median class size in elementary school classes from 30 in 1961 to 24 in 1986. Similarly, the median class size for secondary schools decreased from approximately 27 to 22 from 1961 to 1986. However, what actually takes place in a smaller classroom, 18 students, that cannot occur in a classroom of 28 students? How does having a smaller class of students impact the classroom setting and student learning? Aside from having less students in the classroom, can class size actually impact student achievement? If so, what grade level(s) experience the greatest improvements?

Flake et al (1995) report that class size reduction may not have the expected effect on student achievement as some educators believe. These researchers report that Utah, a state with the highest student/teacher ratio scored in the top ten in SAT, ACT, NAEP, and graduation rates. Conversely, New Jersey, with the nation’s smallest student/teacher ratio, ranked in the top ten in only two areas and number 37 in SAT scores. Despite these findings, several other studies indicate otherwise.

Although class size reduction may improve student achievement, it also entails other factors that tend to be overlooked. The demand for qualified teachers, support staff, and infrastructure, i.e. additional classroom space and additional utilities, stem from class size reduction. In addition, different attempts have been made to initiate class size reduction at various grade levels. However, each of these attempts differs not only in approach but also in study population chosen for their respective studies. With class size reduction's popularity, it becomes important to determine whether the efforts are worth the related costs, both direct and indirect. Direct costs include the actual dollar amounts associated with this initiative, e.g. additional classroom space, teachers, and materials. Indirect costs include such components as hiring inexperienced or unqualified teachers and having to cut school programs to fund this type of initiative. Although direct costs impact school districts financially, indirect costs such as hiring unqualified teachers may impact the actual initiative being implemented. Without qualified teachers to implement a program, the effect may never reach its full potential.

With a great deal of money and effort being associated with class size reduction, accountability must also be taken into consideration. If class size reduction does promote student achievement, then it becomes important to determine what grade levels benefit the most and the cost associated at these grade levels.

### *Statement of the Purpose*

Meta-analytical studies do exist in the educational field. In fact, meta-analytical studies, such as the study conducted by Glass and Smith (1979), reviewed issues such as class size and its effects on teacher-student interaction. Smith and Glass (1980) focused more on teacher attitudes towards teaching in classes with more versus fewer students rather than its impact on student achievement. In addition, Goldstein, Yang, Omar, Turner, and Thompson (2000) report more on the effects of a meta-analytical study in educational research rather than reviewing the actual effects of class size on student achievement.

Although class size reduction studies have been conducted, their findings have varied. In fact, Flake, Von Dohlen, and Gifford (1995) claim there is no strong evidence at the national level that smaller class sizes have improved student performance. However, several studies have taken place in the efforts of reducing class size and examining its effects on student achievement. The main concern is that each of these studies varies in approach. For example, one may only entail one school district while another study may be a statewide effort. Furthermore, research is plentiful enough that, examined one study at a time, a few studies can always be found to buttress whatever position is desired, while the counter-evidence is ignored. Thus, advocates for or against class size reduction often cite isolated studies either to support or refute its value. By using a meta-analytical approach for this study, results can be reviewed despite differences between the data entailed in these initiatives. Variables studied in this investigation include ethnicity, gender, grade level, and content area.

With the fiscal implications associated with mandated class size reduction, both at the state and local taxpayer level, this study becomes worthwhile to determine the most effective

areas for such an initiative. In addition, the purpose of this quantitative study is to provide a meta-analysis on research pertaining to student achievement as related to class size. The inquiry of this study is aimed at determining what mediators have the greatest impact on student achievement.

*Question of the Study*

1. What variables mediate the relationship between class size reduction and student achievement, e.g. ethnicity, gender, grade level, and content area?

This question will be answered by reviewing recent data and information provided from professional research articles and journals. For the purpose of this study, twenty-six studies were compiled and reviewed for the meta-analysis. Three contained sufficient data and were included in the meta-analysis. It is my intent that this compilation of information will assist policymakers in understanding the impact of class size amendment.

### *Definition of Terms*

The following are definitions of key terms used in this research project.

*Class size amendment* is any legislation, state or local, passed that entails reducing student class size to specific limits.

*Content area class* is a class that be classified as reading, math, science, or social studies.

*Effect size* is the degree to which the phenomenon is present in the population or some specific non-zero value in the population (Cooper, 1998). *Effect size* is also how large a difference you want to detect in your study, not in terms of an actual score but in terms of a  $z$  score (Shavelson, 1996).

*ESOL* refers to any student that can be classified as an **English Speaker Of another Language**.

*Ethnicity* is the term used to describe one's origin or background.

*Exceptional education* refers to any student that is classified as having a learning, mental, or emotional disability.

*Fixed effect* is used to describe an effect size when the only random influence on it is sampling error.

*Gender* is the term used to describe the sex of a subject.

*Grade level* is the year in school a subject is in according to a traditional K-12 school setting.

*Minority student* is any student whose racial or ethnic background is other than White non-Hispanic.

*Mediator* is a variable that is analyzed to determine its relationship to any additional variables presented in the study.



*Random effect* is used to describe an effect size in a data that is affected by a large number of uncontrollable influences.

#### *Assumptions and Limitations*

Certain limitations exist in this study. First, sufficient existing studies that represent class size reduction may not be available or current. This lack of information may be the result of assessment data not being recorded. In this case, the data available will be reviewed for inclusion in this meta-analysis. Finally, all of the data collected may not include specific information pertaining to ethnicity, gender, grade level, and content area.

This study contains the following limitations:

1. This study depends heavily on research information pertaining to class size reduction and its effects on student achievement. This information may not include areas related to ethnicity, gender, grade levels, and content area classes as mediators.
2. This study only reports the meta-analysis of the studies collected.
3. The reports collected may not contain ethically and accurately reported data.
4. The reports collected did not describe the details of the assessments administered to students. Therefore, each standardized test may have assessed different skills and concepts.

This study contains the following assumptions:

1. The reports collected contain data that includes exceptional education or ESOL students and not only regular education students.
2. The reports collected contain data that includes an equal number of males, females, white, minority, and grade level students.

3. The reports collected contain data that includes reporting results for students from varying socioeconomic status.

### *Methodology*

#### Research Design

This study is a meta-analysis statistical report on class-size reduction. The following variables were examined within each study and include ethnicity, gender, grade level, and content area, when available.

#### Population

This study examined archival data and did not obtain information from human subjects. The data utilized in this study were collected from various class size initiatives. The data from the class size initiatives contained information pertaining to student achievement based on ethnicity, gender, grade level, and content area. Financial components can be defined as the dollar amount associated with this type of initiative. Differences existed among the various class size initiatives. Different ethnicities, genders, grade levels, and content areas were examined in each initiative.

#### Instruments

After being obtained from educational sources, class size initiative data were analyzed using the *Comprehensive Meta-Analysis* (Biostat, Inc., 2000) software.

#### Data Collection

No human subjects were in the collection of the data used for this study. Using the data collected from the various class size initiative studies, the mean, standard deviation, mean effect size, and student achievement results were analyzed. With the use of meta-analysis, ethnicity,

gender, grade level, and content area could be examined regardless of the age of the study, population size, participants, and geographic location.

The following variables were coded and entered accordingly: ethnicity, gender, grade level, and content area. The data were then analyzed using statistical procedures to determine if significant differences exist between the groups that include the variables listed. In addition, statistical information, i.e. mean, median, and standard deviation, effect size for each class size reduction initiative, when available, were coded and used in the meta-analytical study.

Furthermore, effect sizes can be categorized into one of three general types, 1) the correlation,  $r$ , 2) the standardized mean effect size,  $d$ , or 3) the odds-ratio. For the purpose of this meta-analysis, only studies that contained either the correlation,  $r$ , or the standardized mean effect size,  $d$ , were examined. When necessary, the correlation,  $r$  was converted into the standardized mean effect size,  $d$ .

For computation and input of the correlation,  $r$ , the correlation provided in the report and the sample sizes per condition or the  $p$ -value are needed. For computation of the standardized mean effect size,  $d$ , the following information is reported for the computer software to generate outcomes: 1) means and standard deviation per condition, per sample size, or 2) means and standard deviations and  $p$ -values, or 3) differences in means, common standard deviations and sample size, or 4) standardized mean effect size,  $d$  and sample size, or 5) mean sample size and  $t$ -values ( $t$  being the square root of  $F$  value), or 6) difference in means, sample size, and  $t$ -value, or 7) sample size and  $t$ -value only, or 8) means, sample size, and  $p$ -value, or 9) difference in means, sample size and  $p$ -value, or 10) sample size and  $p$ -value.

Although research studies have been conducted in the efforts of analyzing class size with factors such as teacher-student interaction or teacher attitude as in Glass and Smiths' 1979 and

1980 studies, other studies have also been conducted to review the relationship between class size and student achievement. In addition, researchers have debated the actual effects of class size on student achievement.

## CHAPTER TWO: REVIEW OF LITERATURE

### *Class size reduction and Educational Policy*

Educational policy to improve student achievement has been the aim of schools for many years. Policy makers have explored various options including staff development, rezoning, alternative assessment, and magnet programs. While these options have been met with varying degrees of success, another option has gained attention as a means to improve student achievement. Legislative and public attention regarding class size as an option for improving student performance in schools has increased.

Although class size may appear to be a newer term in the educational field, it has long been revered as being a positive influence on student achievement and other teaching components. According to Hallinan and Sorensen (1985), in a 1974 NEA teacher opinion poll, 80 percent of the teachers polled believed class size was “extremely important” in improving student achievement, whereas the remaining 20 percent considered class size to be “moderately important.” Furthermore, Hallinan and Sorensen report that in a 1975 teacher poll, lowering class size was named more than any other item as the one improvement that would create better teacher morale and job satisfaction.

Due to the attention given to class size, it has found its way into many different areas in the educational field. With class size gaining public attention, legislative initiatives have also followed suit in the efforts of improving school performance. Despite the popularity of smaller class size, factors associated with class size have proved inconsistent. After reviewing these factors, do smaller class sizes lead to educational benefits for students? In recent years, the standards-based movement has led to a call for accountability in public schools. This in turn has

only promoted the implementation of smaller classes. Although in theory this may in fact promote student achievement, costs have been overlooked. Not only is there a direct financial impact imposed by this type of initiative, this project may also entail indirect costs such as hiring unqualified teachers. This alone may have dramatic impacts on classroom learning and teaching. Conversely, smaller class size may still prove beneficial to student achievement regardless of teacher ability and efficacy.

States across the country have implemented class size reduction programs with the hopes of proving its positive impact on student achievement. Tennessee, California, Wisconsin, and Arizona all implemented state-funded initiatives that targeted smaller class sizes. However, each of these programs included different experimental groups and varied in how long the program existence.

Financial components associated with class size reduction range from a few hundred thousand dollars for smaller school districts to hundreds of millions for larger school districts. School districts have now had to view money allocation differently to accommodate such programs. This effort can make it difficult for school districts to achieve both financially and structurally.

Legislative influence has grown to support the need for class size reduction. Rice (1999) states that many policymakers continue to view class size reduction as a promising policy for improving student achievement. Rice indicates that legislators argue smaller classes have the potential to improve student achievement. Pong and Pallas (2001) state the reforms necessary for improving student achievement will only take place when teachers have fewer students. Pong and Pallas report that in the United States a great deal of evidence has shown that smaller classes can improve student achievement during early childhood schooling. The U.S. Department of

Education (2000) reported in Washington D.C., first grade teachers at Hendley Elementary School recount having greater satisfaction with students' achievement, motivation, and skills when they are able to provide instruction to a smaller number of children. The Department of Education report further states the National Education Act of 1969 lists class size reduction as a recommendation for school improvement. According to Pong and Pallas (2001), the consideration for improving students' achievement by reducing class size has remained an option in the United States longer than in other countries due to this act.

### *Defining Class Size*

Many educators refer to class size reduction as the source of improving student achievement. Odden (1990) indicates measuring class size is not straightforward. Merely using a pupil-professional ratio that includes certified staff is not accurate because certified staff include both professionals that teach in the classroom and those who do not. In addition, the "average" pupil-teacher ratio can be misleading because many teachers have non-classroom duties. For the most accurate definition of class size, education researchers need to examine the actual number of students in the classroom. Despite this claim, several studies have composed their efforts around various definitions of class size. Goldstein and Blatchford (1998) state that defining class size is very problematic. These researchers indicate true class size varies from day to day. For example, students listed on a class roster differs from the actual number of students being taught in class on any given day. In an elementary class, as Goldstein and Blatchford point out, students may be removed from a class for various reasons, thereby changing the class size. Therefore, to rely on a formal class size, i.e. counting at the start of the year, may prove to be a very poor guide to actual student experiences.



Brewer, Krop, Gill, and Reichardt (1999) report the term “class” can be defined differently across policies. Some states may use student-teacher ratio, some states use actual class size, while still other states allow teacher aides to be included in the calculation of class size. According to Flake, VonDohlen, and Gifford (1995), class size can be expressed in two ways: student/teacher ratio and students per class. Although various articles and studies interchange these terms, these two can easily be connected. Flake et al state that generally a student/teacher ratio is at least 10 students smaller than the typical class size. Therefore, a teacher/student ratio of 15:1 indicates an average class size of at least 25. Students per class tends to provide a more accurate depiction of class size. However, for this study, both terms will be used depending on the source from which it was retrieved.

#### *Past Class Size Reduction Study Results*

Before analyzing results of class size reduction can take place, a look into the factors associated with class size reduction must first be reviewed. The effects of class size on student achievement have long been a discussion topic for educators to debate. Glass and Smith (1979) claim that the research around smaller class size and student achievement is unclear. Glass and Smith state teachers spend a great deal of time thinking about class size and administrators insist on increasing class sizes due to the economic importance associated with this issue. The researchers point out teachers focus on how smaller class size will benefit their classes in such areas as classroom management, student interaction, more time for in-depth curriculum. School administrators review class size as a financial burden because of the added costs required to hire additional staff, provide classroom space, and materials.

Although class size appears to be over-researched, these authors make it clear that much can still be learned about class size and its effect on student achievement. Goldstein and

Blatchford (1998) claim that despite the number of reviews of class size studies, there appears to be no consensus about different size classes promoting student learning. Good, Grouws, Mason, Slavings, and Cramer (1990) state small group instruction has shown little evidence in improving higher order mathematical thinking or problem solving skills. Good et al indicate that little research is available that tests the effects of small group processes. Therefore, further research is needed to determine the true effects of class size on student achievement.

As Goldstein, Yang, Omar, Turner, and Thompson (2000) state the effects of class size on student achievement have been studied both qualitatively and quantitatively for several decades. A large number of studies pertaining to class size exist. These studies include observational surveys, matched designs, and randomized controlled designs. Although each of these studies stated improved student outcomes, none of them have produced conclusive results linking the two topics together.

#### *Conflicting Arguments with Class Size Reduction*

Educators have debated the exact teacher/student ratio that produces the highest student achievement. Unfortunately, one study may draw one conclusion while another study may claim a completely different outcome altogether. In addition, with no official nationwide program aimed at reducing class size, individual states and local school districts have been left to discover this initiative on their own. Also, studies that include class size reduction have varied in nature. According to Finn and Achilles (1999), few definitive conclusions can be drawn from previous class size research because some studies employed nonexperimental designs. Furthermore, many of the studies were small samples or were of short duration. Pong and Pallas (2001) state the concerns associated with class size reduction are related to student assignment into a class and may depend on students' prior academic performance. If higher ability students are placed in a

small class, the greater the possibility the class will perform well when as compared to lower ability students will result with the same teacher-student ratio. This may imply the class' performance is based on students' abilities prior to being part of the class rather than the actual class size itself. Additionally, if high socioeconomic (SES) students move to schools with small class sizes, their higher achievement level may be an effect due to economic background and not small class size. The same can be concluded of lower-achieving students placed in smaller "remedial classes".

#### *Universities and Class Size Reduction*

Although this research will focus primarily on public K – 12 education it is still important to explain class size reduction and its impacts on other levels of education. For example, state colleges and universities, are still under legislative control. Therefore, class sizes and the financial impacts on universities should be examined. In a study, Maxwell and Lopus (1995) classify small class size as no more than 55 and a large class of no more than 120 students. With these numbers in mind, grades K – 12 cannot even be linked to university classes. However, Maxwell and Lopus point out the financial benefits of having larger classes in the university as compared to smaller classes. For instance, they note the cost of educating 132 students in one large class is \$12, 551, or an average of \$95.08 per student. In comparison, educating the same number of students in three smaller classes is \$17, 632 or \$133.58 per student. In their study, Maxwell and Lopus reviewed the performance of college students in an economics class. Although there was no evidence that proved the students learned more in a smaller class setting, according to the multiple-choice tests used in the study, there was proof that students no longer enrolled in economics when they were part of a large class.

It is important to note that a great deal of money, this will be discussed later in this research collection, and effort has been placed into class size reduction in the public school system, notably grades K – 12. Furthermore, many studies have been recorded to present their results about class size reduction and its effects on student achievement. Despite this, each study has differed in some form or statistical collection. Therefore, no real connection between these studies has been made.

As parents or voters call for smaller class size, politicians continue to impose these efforts on school districts. Typically following suit is high accountability standards for school districts to meet. Legislative accountability measures have been placed on educators and students in contemporary years. Recent legislation has led to an accountability movement that requires students to achieve at expected levels of knowledge within a specific amount of time. Standards-based learning has been the focus of legislators and educational policy makers.

#### *Standards-Based Accountability and Class Size Reduction*

With the standards based accountability movement in educational reform, where does class size fit in? So much has been added to education reform in recent years, it can become rather confusing as to how class size reduction will truly impact student performance. According to Flake et al, Title 15 of Arizona's education code originated with only 242 pages in 1956 but has grown to 1,119 pages in 1995. Arizona's Title 15 is an 18 chapter portion of the state's Constitution, that includes the governing educational laws school districts must follow. It appears that class size reduction has taken a step in the mandated direction. Rather than legislation focusing on "what is best for students to learn", i.e. standards, they have focused more on "We know what size classes your child be placed in."

With the amount of federal dollars, as well as state and local funds, it becomes vital to examine the true purpose of class size reduction. Legislators and policymakers will fail to see the need to allocate federal dollars to class size reduction if positive results do not follow. To review the research, class size reduction programs need to be analyzed to determine where the most impact is made and if this effort is truly worth the cost.

Accountability standards impact how student achievement is perceived by different stakeholders. Policymakers need to be ensured that student achievement gains result from federal dollars. Parents hope their children attend high-performing schools. Thus, low-performing schools have become a focus of legislative efforts. Low-performing school districts, like Columbus, Ohio, are now able to use federal funds to provide professional development and support needed for effective implementation of smaller class size models.

Krueger (2001) indicates improved school resources like smaller class sizes increase test scores, which, in turn, predict labor market earnings. Krueger claims the economic benefits in class size are at least equal to the costs. Krueger states a strong relationship between student performance and future earnings exists due to a reduction in class size. With this in mind, it becomes easy to see why so many government funds are being placed into class size reduction programs and why there is an economic benefit behind this effort.

#### *Direct and Indirect Costs Associated with Class Size Reduction*

Despite the costs associated with class size reduction, both direct and indirect, many educators still believe the benefits far exceed these costs. Direct costs include the financial components associated with class size reduction. Money is needed to build classrooms, purchase materials, utilities, and pay teacher salaries. These can be categorized as the direct costs associated with class size reduction. Indirect costs are the offspring of direct costs. Indirect costs

include hiring inexperienced or unqualified teachers. This practice can have negative impacts on student achievement based on teacher inexperience. Inexperienced teachers may bring a lack of classroom management, instructional practices, curriculum knowledge, and interaction skills.

#### *Effects on Classroom Instruction*

Research has been conducted to determine whether student achievement can be justified by the amount of money being spent. Although some educational researchers may claim no such gains in student achievement exist with class size reduction, an almost equal amount of researchers indicate otherwise.

Villaire (n.d.) states that smaller classes enable more one-on-one contact with the teacher, additional written and oral feedback for the students became available, and more opportunities for the students to ask questions and respond during class followed. Furthermore, as teachers become more acquainted with the students, students become more acquainted with each other. Karen Schwengel, a former teacher, as stated by Villaire, indicated being able to focus on teaching instead of classroom management had a powerful effect on teacher morale in her school. Villaire also said the results are clear in smaller classes and that teachers can become more attentive to individual students, the classroom atmosphere is more cohesive, behavioral problems decrease, and teacher morale increases. In addition, fewer students drop out of school and fewer are retained.

Bourke (1986) analyzed different teaching practices and class size to determine whether statistical significance existed between them. Bourke reviewed the following teaching practice variables: 1) use of whole class teaching, 2) number of groups used within the class, 3) any interactions between teachers and students, 4) amount of time tolerated in class, 5) non-academic management, 6) teacher probing after a question, 7) student questions, 8) teacher waits or a

response, 9) homework and assignments assessed, 10) teacher directly interacting with students, 11) teacher monitoring student work, 12) positive teacher response to answer from student, 13) teacher lecture versus explanation, and 14) use of oral tests for assignment. Bourke concluded none of the individual teaching practices proved significantly statistic with student achievement. Rather, when the teaching practice variables were combined with class size, a significant correlation with achievement became more apparent. Therefore, for this study, it can be postulated that class size can have a positive impact on student achievement.

Bourke (1986) further discussed two major aspects stemming from class size reduction. These two included the nature of the teaching practice variables found to be related to class size and student achievement and the nature of the relationships found in the causal model, taken as a whole. More specifically, in smaller classes, the teachers taught the class as a whole rather than in small groups and yet still displayed higher achievement. On the contrary, teachers with larger classes tended to form small groups that led to lower achievement levels. Bourke points out this difference can be attributed to the teacher having to instruct and explain the material repeatedly to different groups of students. Rice (1999) reports on teacher's use of time in a smaller class setting. For instance, Rice critiques whether a teacher would actually change his/her teaching methods due to having a smaller class size. Imagine if a class was reduced from 30 to 20 students. Would the teacher actually take the extra time to learn and teach additional instructional methods aimed at improving student achievement? If not, the teacher could continue teaching the same exact way, using the same exact methods, as was done in the larger class of 30. This may be, according to Rice, voluntary or because the teacher is not knowledgeable in teaching methods useful for smaller classes.

### *Effects on Student and Teacher Interaction*

Bourke (1986) reported smaller classes had fewer interactions between the student and teacher despite higher achievement being recorded. More specifically, two types of interaction were found to differ significantly. First, student questions to teachers were more frequently for clarification in larger classes than in smaller classes. Despite this, there was more teacher follow-up of questions in smaller classes. Smaller classes also provided additional time to focus on items such as homework, class assignments, and other assessments such as oral tests. Second, larger classes typically required more management interventions than did smaller classes thereby reducing the amount of instructional time. Bourke continued to report that teachers in smaller classes probed their students more frequently with effective questioning techniques that led to higher achievement. When less time is spent on management, more time can be spent on probing and wait time both of which, according to Bourke, are more productive uses of class lesson time. With these two components incorporated in smaller classes, student achievement improved. In addition, more of a lecture approach was used in larger classes. Although this increased teacher time on instructional delivery, this did not necessarily improve teacher-student interaction.

Bourke (1986) also indicated homework and noise level were impacted by class size. Smaller classes were given more homework than their counterparts and showed improved achievement. Homework tends to be a more productive activity. With smaller classes facilitating the setting, more time could be spent on homework correction. In addition, as might be expected, in this study, smaller classes were not as noisy as larger classes and had higher achievement. Therefore, a claim can be made that smaller class size improved student achievement. In larger classes, more interaction exists overall between teachers and students that leads to higher noise levels and additional management concerns.



Betts and Shkolnik (1999) report the shift in teacher time being caused by changes in class size. More specifically, their results indicated a significant effect on teacher time variables. Teachers with smaller classes had more time to devote to instruction due to less time being spent on discipline. In addition, more time could be devoted to individualized instruction rather than whole-class lecture situations.

Therefore, Betts and Shkolnik (1999) concluded class size increases impact three major types of class time. First, class size increases led to more time being spent on administrative duties, i.e. classroom management. Second, the amount of time teachers spent reviewing in class also decreased as class size increased. Third, Betts and Shkolnik recorded no significant effects were discovered for time spent on new material or percentage of textbook curriculum presented. With smaller classes, teachers of lower scoring classes and middle school classes, exhibited a shift from lectures towards individual instruction and increased instructional time. Also, in smaller classes, teachers strived to improve student understanding by spending more time on review rather than addressing new topics.

Betts and Shkolnik (1999) explain their findings of instructional time using the following example. Imagine class size being increased to 40 from 20 students. Based on their study, the two largest predication shifts in class time are a 3.0% decrease in review time and a 2.5% increase in classroom management time, which includes student discipline. In a typical class that meets for 300 minutes a week, this computes to nine minutes less on review and seven and a half minutes more on discipline. Although this may seem negligible, consider this over a 36-week period. Both numbers now increase to 324 minutes of review lost and a total of 273 minutes spent on discipline. This is approximately a week of instructional time lost.

The independent Educational Priorities Panel completed a study in 2000 for the class size reduction program in New York City. According to Haimson (2000), class size reduction results in the following improvements: 1) noticeable declines in discipline referrals, 2) improved teacher morale, 3) an increased focus on prevention rather than remediation, and 4) higher levels in classroom participation by students. To support this, then U.S. Secretary of Education Riley stated, “Smaller class sizes allow our teachers and students to do the best they can. Teachers do not teach most effectively when they are hampered by the burden of too many students in the classroom” (U.S. Department of Education, 2000, p. 8).

### *Class Size Reduction and Student Achievement*

The impact that class reduction has on student achievement based on class size still needs to be examined. Smith and Glass (1980) studied the effects of performance when students moved from different classes with different class sizes. Consider an average student in a class of 30 being set equal to the 50<sup>th</sup> percentile. According to the authors' study, if this student were placed in a class of 20 students, he or she would experience benefits higher than 58% of his or her peers in a class of 30. Now consider if the student were placed in a class of 10 students. One could expect, as Smith and Glass report, an increase to the 70<sup>th</sup> percentile of his/her original class of 30. Conversely, if the student were placed in a class of 40, the student would fall to the 45<sup>th</sup> percentile when compared to his/her peers in the class of 30. Hypothetically, the student would fall further down to the 39<sup>th</sup> percentile if placed in a class of 60 students. Levin (1988) notes increases in student achievement due to class size reductions. Levin points out the most notable student achievement occurred when classes in reading and mathematics were reduced from 30 to 15 students.

According to the U.S. Department of Education (2000) report of the class size reduction program, small classes were stated to promote stronger bonds between teachers and students that led to an improved school climate and fewer discipline concerns. In addition, teacher morale improved due to the increase of time spent on classroom teaching rather than classroom management. Flake et al (1995) state every study researched in class reduction improved teacher morale. In the Rockford, Illinois School District 205, teachers claimed they were able to increase the attention given to students and this improved early identification of reading problems as a result of the smaller classes.

Raising reading achievement levels has most likely been the focal point of all class size reduction efforts. At Georgetown Elementary School in Anne Arundel County, Maryland, reading scores improved in first and second grades to the point the school is now one of the top three elementary schools in the county. Furthermore, in West Middlesex, Pennsylvania, as a result of reducing class size in two elementaries, during the 1999-2000 school year, student scores increased in reading, writing, and mathematics on the Iowa Test of Basic Skills. The students' overall grade equivalent scores increased from 1.9 to 2.1. The Trinity Area School District in Pennsylvania, between 1999 and 2000, showed a growth of 12 percentage points in the number of first grade students scoring at 80 percent or higher. Similar results were recorded in their second graders with an increase of 21 percentage points (U.S. Department of Education, 2000). Swenson (1991) reports smaller class sizes are necessary for teaching language to students. Larger class sizes restrict interaction time, both formal and informal, between teachers and students. Swenson further claims, "if we believe in the centrality of language,... then we must have smaller classes in which to foster that language development (p. 13).

On the other hand, Pong and Pallas (2001) describe how smaller class sizes, when compared to Asian countries, had little effect on student achievement. Interestingly, class sizes in both Japan and China, especially in the elite classes, were larger than those in the United States yet they exhibited higher student achievement. The researchers also pointed out that ethnographic studies uncovered Japanese and Chinese teachers viewed little relationship between class size and student achievement. However, it appears in the United States, much attention is still placed on small class size.

### *State Class Reduction Programs*

In 2000, the National Government began to view the positive results of class size reduction funds allocated to states and individual school districts. Over 20 states, including Georgia and Massachusetts, had incorporated class size reduction programs within their school districts in the efforts of improving student achievement. While some states' programs are newer than others, all still focus on the efforts of improving student achievement in all academic areas, especially reading, writing, and math. Waymack and Drury (1999) report significant test score gains for students enrolled in small classes, across all subject areas.

#### *Tennessee's Project STAR*

Tennessee's Project STAR (Student Teacher Achievement Ratio), established in 1985, was among the first and probably most well-known class size initiative. Although this program paved the way for many programs to follow, it lacked credibility for many reasons. First of all, STAR was more of a case study than an amendment implementation. More specifically, Gilman and Kiger (2003) make note how STAR only included 79 schools from rural, suburban, urban, and inner-city settings. The only requirement for the participating schools was that they have a large enough population to have a control group and two treatment groups at any grade level. Kennedy (2003) describes how STAR entailed reducing class size to no more than 13 – 17 students from classes with 22 – 25 students.

According to the North Carolina Board of Education Evaluation Brief (2000), STAR involved randomly assigning K – 3 students to one of three class configurations: small class (S) with class sizes of 13 to 17, regular class (R) with class sizes of 22 to 27, and regular class with a full-time teacher aide (RA) with class sizes of 22 to 26. Teachers were randomly assigned to the

classes. Students remained in the same class until they reached the third grade. Once they reached the fourth grade, they were placed into regular classes.

Gilman and Kiger (2003) report four research findings from the STAR project. First, students in smaller classes performed better on all sets of achievement measures. Secondly, positive achievement occurred despite school location and gender. Third, some minority students or students attending inner-city schools received greater benefits of student-teacher interaction and success in school. Finally, there were no differences in small and larger classes in student scores on motivational scales. Florida Tax Watch (2004) further reported that STAR impacted kindergarten students the most. In addition, Weglinsky (2004) indicates achievement differences continued through the eighth grade, four years after students in small classes had been placed in regular classes. Furthermore, achievement gains were more apparent for minority students than white students.

Follow-up studies were conducted to STAR in 1989 to determine whether the study contained any long-term effects. The North Carolina Board of Education Evaluation Brief (2000) reported students classified in three different classrooms. (R) classes of regular size where no reduction took place. (S) classes of smaller size where reduction was implemented. (RA) of regular size however contained a teacher aide as well as the teacher. This report states students who had been in (S) classes during grades K – 3 experienced higher achievement level in all academic areas as compared to students who had been in larger (R) and (RA) classes. In addition, students in (S) classes were better behaved than students from larger classes. Finally, students from (S) classes were rated as spending more effort time in class while taking more initiative in their learning.

According to Armentano (2003), there have been nearly 300 studies conducted nationwide targeted at reviewing class size and student achievement. Armentano continues to report the following:

Professor Eric Hanushek, an economist at the University of Rochester, reviewed these studies and discovered that only 15% of studies suggest that reducing class size improves student learning as measured by standardized tests. Indeed, in 72% of the studies reviewed, there was no statistically significant effect on measurable student achievement associated with smaller classes. Even more surprisingly, in 13% of the studies reviewed, student test scores actually declined as class size was reduced. In sum, a full 85% of all of the studies on class size and student achievement found that reducing class did not improve student performance (p. 1).

Armentano (2003) states class sizes have been falling nationwide for several decades. This is partly due to the fact that several states and districts began to initiate class size reduction programs. For example, class average was 30 students per class in 1961 and only 23 in 1998. Despite this, there has been no substantial improvement in student classroom performance when measured by standardized tests.

#### *California's Class Size Reduction (CSR)*

Chudrewicz (2002) reports some growth in student achievement due to California's Class Size Reduction (CSR). More specifically, the most positive result of the CSR experiment is a small increase in third grade achievement. In addition, the North Carolina State Board of Education Evaluation Brief (2000) report that along with improved results in third grade students, for some students who moved to fourth grade in 1998-99, a small positive effect also exists, even after returning to large classes. Despite this, more damage was done for other students, namely minority or exceptional education students. Due to the decline in teacher qualifications, ESOL and minority students under-performed on standardized tests. Chudrewicz adds that about 1000 teachers changed to teaching kindergarten through third grade instead of teaching ESOL and exceptional education students.

In a time when educational funds have dwindled and the public has been searching for the appropriate use of these funds to ensure student achievement, added pressure has been placed on educators to review costly educational programs. Kennedy (2003) reports that only inconclusive evidence can be found when linking smaller class size to student achievement. Kennedy states the Class Size Reduction Research Consortium 2002 report could not show that class size actually improved student achievement. Although test scores have improved since the start of the class-size reduction program in 1997, no direct link has been given to class size. Gilman and Kiger (2003) further support this claim by indicating that California's (CSR) does indicate gains but they cannot be directly attributed to class size reduction. Gilman and Kiger state that CSR resulted in a reduction of elective-type courses and an increase in core curriculum classes. Therefore, more academic focus had been placed during the school day and may have had an impact on student achievement greater than small class size.

#### Wisconsin's SAGE

Wisconsin's class size initiative, Student Achievement Guarantee Education (SAGE), focused on grades K-3. The University of Wisconsin – Milwaukee, according to Kennedy (2003), concluded the major difference in smaller classes (SAGE) is increased individualization. The North Carolina State Board of Education Evaluation Brief (2000) indicates the target class sizes of SAGE included 15 for grades K-1 in 1996-1997, grade 2 in 1997-98, and grade 3 in 1998-99. Districts with a high number of low-income families were selected as eligible for this program.

SAGE schools demonstrated statistically higher scores on the Comprehensive Test of Basic Skills (CTBS) in language arts, math, and total scores than comparison schools. Smith, Molnar, and Zahorik (2003) used average growth curves, a means of charting and predicting the



average performance of a test's norm group over time, to measure SAGE's effect on the achievement of the cohort of SAGE students entering the 1<sup>st</sup> grade in 1997-1998. Smith et al reported the average growth curves for the CTBS tests used in the SAGE evaluation indicate students would be expected to gain, on average, 33 points in reading and 29 points in mathematics between the fall and spring of first grade. While both SAGE and comparison schools performed better in the spring administering of the CTBS, SAGE schools gained 47 points in reading and 42 points in mathematics while the comparison schools' students only gained 38 points in reading and 35 points in mathematics.

After adjusting for individual pre-test results, socioeconomic status (SES) and attendance, the North Carolina State Board of Education Evaluation Brief (2000) reports small classrooms tended to score higher in language arts, math, reading, and total scores when reviewing scores at the classroom level. In addition, during both years, although scoring lower on CTBS pre-tests, African-American students in smaller classrooms made larger achievement gains on CTBS post-tests than African-American students in comparison schools. More specifically, Smith et al (2003) report African-American first grade students outscored the comparison schools by 14 points. This can be translated into a two-thirds of a year academic growth.

African-American students, according to the North Carolina State Board of Education Evaluation Brief (2000), closed the achievement gap with White students by making greater gains on the total scale score of the CTBS. In comparison schools, African-American students made smaller gains when compared to White students. SAGE reports indicate results at the end of the first grade continued into the second and third grades. Smith et al (2003) further support this claim by indicating SAGE African-American second and third grade students kept pace with their White peers in student achievement. Furthermore, when compared to students in

comparison classrooms, the achievement gap between African-American and White students dramatically decreases in first grade and is prevented from increasing in second and third grades.

Smith et al (2003) continue to report the same general pattern in CTBS mathematics results. Their study shows African-American first grade students made a 10-point gain between pre-test and post-tests. This increase entailed a 19-point gain when contrasted with students from comparison classes. When reviewed with the average growth curve, a means of charting and predicting the average performance of a test's norm group over time, this also represented a two-thirds of a year's growth. This growth continued into the third grade year with an overall 29-point gain, representing nearly a full year's growth over African-American students in comparison classrooms.

Despite these results, Florida Tax Watch (2004) reported SAGE produced statistically higher student performance in language arts and math in first, second, and third grades. However, this report also indicated SAGE concluded different results under other conditions. For example, this report stated that smaller classes in the second and third grades had minimal impact and, in some instances, no additional impact on student achievement. In addition, the actual long-term effects for minority students will be unclear even though these students made gains in first grade. Furthermore, SAGE classes had differing effects on achievement, depending on the subject area. Smith et al (2003) added to this argument by stating SAGE had differing results with students with varying socioeconomic status (SES). Initially, students with low SES outscored students of higher SES by seven points. However, by the end of the third grade, students with lower SES fell behind by 15 points, thereby indicating a 22-point swing. In summary, although SAGE produced results for all participating students, African-American students, especially those of higher SES, benefited the most.

### Arizona's Class Size Reduction

Flake et al (1995) conducted a study of Arizona's class size reduction program from 1992-1993. This report studied the prospect of improving student achievement by lowering class size in grades kindergarten through third. More specifically, the authors report the mean student/teacher ratio in this study was 19:1 for elementary level classes and 16:1 for secondary levels. In addition, the authors studied the effects of enrollment, expenditures per student, student-to-administrator ratios, and of course, class size.

In their findings, the researchers discovered almost no relationship between class size reduction and average student performance. Also, there appeared to be no relationship between class size and average performance at the kindergarten through third grade levels. For high school students, despite smaller class sizes, Scholastic Aptitude Test (SAT) scores decreased. Flake et al note this decrease in aptitude test scores does not suggest that smaller class sizes cause lower test scores. However, the authors state that if there is a relationship between smaller class sizes and student achievement, it is too weak to overlook other factors associated with improving student achievement.

### Differences in State Class Size Reduction Programs

Although each of the studies presented benefits on student achievement at various grade levels, in various subject areas, we cannot overlook the pitfalls associated with the studies. According to Hanushek (1998), between 20 and 30 percent of STAR's students quit the project each year, resulting in less than half of the original students remaining at the end. Secondly, the students that did not complete the program tended to be lower-achieving students, thereby allowing for improved student achievement in the smaller classes. Third, if a student entered a class once the school year started, no pre-test of any kind was administered. This provided no

benchmark in assessing the student's level of achievement. Finally, the schools or teachers chosen for STAR were not randomly selected. Only the students were randomly placed in the classes.

With each of these studies claiming benefits for their own students, it is interesting to examine how different each study actually was to each other. California's CSR contained several differences when compared to Tennessee's STAR. Chuderewicz (2002) describes the main difference between CSR and STAR to be the fact that CSR was a statewide initiative whereas STAR included only about 10,000 students. In addition, STAR reduced class sizes to between 13 – 17 students while CSR reduced class sizes to no more than 20 students. Furthermore, California, as a state, encompassed a much more ethnically and linguistically diverse population group. Students were no exception to this. With California's growing population, the state also experienced two situations Tennessee did not – inadequate space and not enough qualified teachers. With this being said, this is proof that CSR should be viewed individually and cannot be compared to STAR.

#### *Financial Components Associated with Class Size Reduction*

Reducing class size requires school districts to provide additional classroom space. Along with this additional space comes the need for additional teachers, utilities, and support staff. When President Clinton's Class Size Reduction initiative of 1998 is examined, a great deal of costs can be uncovered. For instance, Brewer, Krop, Gill, and Reichardt (1999) report for grades 1 – 3 to have class sizes of 20, 18, and 15 respectively, as proposed by President Clinton, some 42,725, 102,687, and 226,910 additional classrooms were needed in the 1998-99 school year. The total number of classes in the United States in grades 1 – 3, during the 1997 – 98, was 510,000. This initiative almost doubled the total number of classes needed for grades 1 – 3. This number

would increase by roughly the same amount each year until the 2007 – 2008 school year. In addition, President Clinton’s proposal included the need for an increase in per student expenditures or operational costs. For a class size of 20, 18, and 15, the per student operational cost was \$189, \$448, and \$981, respectively. This cost figure does not include annual classroom operating costs. When combined together, these costs imply a need for an average operational cost of \$562 per student by the 2007 – 2008 school year. Brewer et al point out the smaller the class size, the more expensive each child becomes to the school.

Flake et al (1995) state that lowering class sizes is an expensive proposition. In their study of Arizona’s class size reduction program, they calculated the percentage change in class size to be the reduction of students in a class, e.g. 25 to 15 students as 40 percent change, and adding this same percentage to per pupil expenditures for those students affected. In other words, there would be a 40 percent increase in per student expenditure for this decrease in class size to take place. Therefore, in their study, Arizona would need an additional \$457,644,600 to implement a 15 student class size mandate in grades K-3 alone. In addition, some 6,861 new classrooms would be needed. This, of course, would require the same number of new teachers. With the average starting teacher salary, in 1991, at \$22,179 and with benefits that account for an additional 25% of the gross pay, the total cost of new teachers would \$182,604,143. This does not even take into account the annual increases due to teacher salary raises. Additional factors affecting operating costs exist, according to Brewer et al (1999). For example, newly hired teachers may vary in terms of experience level thereby causing a variation in teacher salary. Teachers may be newly certified, retired and reentering the profession, or current teachers moving from one grade to another grade or location. In addition, education level plays a role in factoring teacher costs. Both of these factors have a direct impact on how a teacher progresses

through the established salary schedule. This is in part, as reported by Brewer et al, to the actual initiative in place. For instance, as with California's class size reduction initiative, more emphasis was placed on hiring elementary teachers. Therefore, more newly certified teachers were needed to accommodate the teacher shortage.

According to the Florida Electronic Budget (2003), Amendment Nine (class size amendment) will require the 2004-2005 budget to include a total of \$517,401, 411 increase request to reduce class enrollment to specific sizes. More specifically, \$154,308,649 for grades PreK-3, \$188,376,863 for grades 4-8, and \$174,715,899 for grades 9-12 thereby totaling the nearly \$518 million budget. In addition, for the 2004-2005 school year in Florida, the class size amendment calls for a reduction of two students in core curricular classes at the middle and high school levels. Core classes include math, science, social studies, and language arts or English. In actuality, class size will require a raise in per student funding or basic student allocation (BSA).

Overall, class size reduction for the state of Florida is equivalent to twenty-five percent of the total budget of colleges and universities, fifty-two percent of the budget for community colleges alone, five percent of the budget for public schools, and three percent of the entire state education budget. Florida's Electronic Budget (2004) allocated \$570 million for the 2004-2005 school year class size reduction. This will provide 40,234 new available student seats or desk spaces. Fair (2002) reports at least \$9 billion alone will be needed for construction. In addition to the construction costs, approximately \$2.5 billion a year will be needed to staff and maintain these additional facilities.

With a cost of nearly \$27.5 billion over the next eight years, Amendment Nine places a great deal of stress on the financial resources. For example, the Florida Tax Watch Research

Report (2004) proposes each of the following will need to take place for the class size amendment to be adequately funded:

- Over the eight years, each household would be required to pay \$3,382 in property tax while paying \$1,343 per capita.
- Even if tourism and other out-of-state taxpayers pay 20% of the cost, Floridians would still need to pay \$1,074 per capita and \$2,705 per household in property tax.
- This amendment would require tripling the corporate tax to generate an additional \$3.4 billion annually.
- A personal income tax of 6.68 percent, if included in current state legislation, to equal the class size annual cost.
- A doubling of the four main tax sources combined (gross receipts, corporate income, beverage, and tobacco taxes) would be needed to meet cost demands.
- The gross receipts tax, which raises about \$800 million annually, would need to be more than quadrupled to meet the need of class size amendment (p.3)

Florida is not the only state that has experienced financial concerns with implementing a class size initiative. For example, according to Gilman and Kiger (2003), Indiana's Project Prime Time, the state's class size project, has made it difficult for several school districts to fund this type of program. More specifically, smaller school districts have found it difficult to afford the required teachers' salaries entailed in this initiative. Gilman and Kiger also report that the tiny school district of Shakamak portrays the struggles of many Indiana school districts. According to Prime Time, first grade classes cannot contain more than eighteen students while grades kindergarten, second, and third cannot have more than twenty. Each of the four grade levels in

Shakamak's elementary school has about 75 students. By dividing each grade into three classes, the elementary school falls short of the required class size. Therefore, although the average class size would be 25, this school still would not qualify for the Prime Time project. With no additional assistance given by the state, the Shakamak school district is left to divert funds from other school programs to fund this program.

Although Tennessee's Project STAR has been regarded as the best known class size effort in the United States, the project cannot fully be compared to more recent class size amendments. For example, Project STAR was not a state-wide program and at the time of its inception cost only \$12 million. Furthermore, Project STAR only lasted for a defined period of time. Despite this, the financial aspect of this program can be examined to show relevance to other class size initiatives. For example, the National Education Association (2002) reports Tennessee's student population has increased by five percent since 2000. Therefore, it is unlikely that Tennessee will further pursue a class size policy. Gilman and Kiger (2003) state that Tennessee also expects to lose 25 percent of its veteran teachers over the next five years. This loss, combined with student population growth, only multiplies the number of new teachers needed to proceed with a class size reduction program.

The Tennessee Education Association (2002) describes how Tennessee has struggled over the past several years to keep highly qualified teachers in classrooms. This is partly due to Tennessee's teacher salary measuring approximately \$5,400 less than the national average. The Southern Regional Education Board (2000) reports that nearly four percent of Tennessee's teachers are under qualified and teaching with waivers or permits. Flake et al (1995) further claim that teacher shortages, due to fewer universities having teacher graduates, make it difficult to meet the need. For example, Arizona colleges only graduated about 800 new elementary



teachers during the 1994 school year. This is only about 12% of the actual amount needed in Arizona's public schools. Therefore, in order to initiate a successful class size initiative, Tennessee, as well as other states, must first look at increasing the number of qualified teachers at higher salaries. The key point here is that more money is needed.

Another state to incorporate a class size amendment is California with its Class Size Reduction (CSR) in 1996. Kennedy (2003) states the program entailed about \$1 billion in its first year. In another study, Brewer, Krop, Gill, and Reichardt (1999) report California spent approximately \$1.5 billion on its class size reduction initiative in 1997-1998. This forced school districts to supplement state funds with an unknown amount of additional dollars in order the mandated state requirements entailed in the program. By the year 2002, California spent about \$1.6 billion. Kennedy (2003) indicates that although this may appear to be a large amount of money, it still falls short of the actual costs associated with implementing a class-size reduction program. Brewer et al (1999) support this claim by stating the 1998 proposed budget did not fund the 100% of the costs of CSR. These additional costs have been left to states and local districts to fund through tax revenue.

The CSR consortium, according to Kennedy (2003), said, "Most districts in our statewide sample reported incurring operating costs for (class-size reduction) that exceeded state payments for it...and these funding problems persisted, or even worsened in recent years" (p. 2). Brewer et al (1999) describe the operational costs of CSR to be additional monetary costs associated with new teachers, staff members, and supplies as well as the cost of new buildings, facilities, and capital equipment. The incentives associated with CSR resulted in a deficiency of qualified teachers required to satisfy all new teacher spaces. Due to the high demand of teachers, Kennedy continues to describe how many California school districts were forced to hire teachers without

full credentials, i.e. without teaching certificates or education degrees. The proportion of non-fully qualified K-3 teachers increased from 1.8 percent to 12.5 percent.

Chuderewicz (2002) adds to this argument by stating in CSR's first year of implementation, some 18,400 new teachers needed to be hired. Of the teachers hired that year, approximately half were inexperienced or lacked certification, of which 30 percent lacked credentials, and 21 percent were hired on emergency permits. These emergency permits meant that the new teacher had a college degree and passed the required teaching exam but lacked formal teaching preparation. In addition, adding some 18,000 classrooms meant changing areas such as libraries, music rooms, computer labs, faculty lounges, and even auditoriums into classroom spaces, either temporarily or even permanently.

Class size has gained attention and has found its way into many different areas in the educational field. While gaining public attention, legislative initiatives have also been initiated in the efforts of improving school performance. However, the factors associated with class size reduction still prove inconsistent. After reviewing these factors, do smaller class sizes lead to educational benefits for students?

The standards-based movement has led to accountability in public schools. This in turn has promoted the implementation of smaller classes as the call for improved student achievement. Although in theory this may in fact promote student achievement, costs associated can be excessive. Not only is there a direct financial impact imposed by this type of initiative, this project may also entail indirect costs such as hiring unqualified teachers. This factor alone may have dramatic impacts on classroom learning and teaching. Conversely, smaller class size may still prove beneficial to student achievement regardless of teacher ability and efficacy.

States across the country have implemented class size reduction programs with the hopes of proving its positive impact on student achievement. Tennessee, California, Wisconsin, and Arizona all implemented state-funded initiatives that targeted smaller class sizes. However, each of these programs included different components such as whether it was statewide or district inclusive. In addition, the lifespan of each program differed.

Financial components associated with class size reduction range from a few hundred thousand dollars for smaller school districts to hundreds of millions for larger school districts. School districts have now had to view money allocation differently to accommodate such programs. This effort can make it difficult for school districts to achieve both financially and structurally.

#### *Summary*

In summary, several issues arise when class size is examined. Policymakers report small class improves student academic achievement while also improving teacher-student interaction. Defining class size is difficult pending the educational environment, i.e. legislators, school districts, and even schools that surrounds its initiative. The actual student benefits from small class size differ due to varying educational reports' purposes and outcomes. In addition, class size varies among the different levels of education, i.e. elementary, middle, and high schools, and even into universities. Standards-based accountability contributes to the support of small class size by indicating higher achievement for students. States have begun to realize a great deal of funding is required not only to implement but also maintain such an initiative. In reality, indirect costs, i.e. hiring unqualified teachers and staff members also have impacts on schools and students. Tennessee, California, Wisconsin, Arizona, and Florida have all attempted to implement small class size programs either experimentally with one district or as a statewide

effort. All have reported tens of millions of dollars being needed to implement and support small class size programs. Despite all these factors, policymakers continue to implement small class size initiatives in the efforts of improving student academic achievement.

## CHAPTER THREE: METHODOLOGY

### *Introduction*

The purpose of this chapter is to report the methodology employed in this study. A meta-analytic approach was used in efforts of synthesizing available data into a generalized form. Ethnicity, gender, grade level, and content area were used as mediators between class size and student achievement. No human subjects were reported in this study as the data was compiled from existing educational studies. Educational reports were reviewed using PsycInfo (2005) and JSTOR (2005) Internet databases. Reports were selected for the meta-analysis based on the data reported. Once reviewed, three educational reports were selected for the meta-analysis. These reports were analyzed using the *Comprehensive Meta-Analysis for Windows* software from Biostat (2000). The results of the study were based on ethnicity, gender, grade level and content acting as mediators between class size and student achievement.

### *Research Design*

This study was implemented following a meta-analysis statistical approach on class size. Ethnicity, gender, grade level, and content area were the mediators reviewed in this study.

## *Subjects*

Due to the nature of this meta-analytic study, no human participants were required for data collection. More specifically, only results of selected class size reduction programs were analyzed. The data utilized in this study were collected from educational sources that documented results for class size reduction as they pertained to ethnicity, gender, grade level, and content area associated with class size reduction. On-line education resources PsycInfo (2005) and JSTOR (2005) were used to collect professional education journal studies. The terms class size reduction and student achievement were researched. Hundreds of on-line articles and journals were retrieved from this search. The difficulty with an on-line search is that each of these terms is searched individually and may not ever be part of a journal or article that also includes the other terms searched. The result is a collection of articles and journals that may only address class but not size, size but not student, or student and not achievement. Furthermore, not all of the articles and journals retrieved in these searches contained hard data or numbers pertaining to the study. After reviewing hundreds of articles extensively, 26 of them contained hard data or numbers presented in a readable manner.

Studies were then selected for inclusion into the meta-analysis based on the information presented. The criteria required for selection into this meta-analytical study were as follows: mean, median, and standard deviation, effect size for each class size reduction initiative. For the purpose of this meta-analysis, only studies that contained either the correlation,  $r$ , or the standardized mean effect size,  $d$ , were examined. When necessary, the correlation,  $r$  was converted into the standardized mean effect size,  $d$ .

## *Procedures*

### *Instrument*

Class size reduction data were analyzed using the *Comprehensive Meta-Analysis for Windows* (Biostat, Inc., 2000).

### Data Collection

Twenty-six studies were collected and reviewed for this meta-analytic study. Three studies collected were analyzed with the meta-analytic procedure. The following variables were examined as mediators between class size and student achievement: ethnicity, gender, grade level, and content area. The data were then analyzed using statistical procedures to determine if significant effects existed between the variables listed. These studies contained data describing class size and student achievement and their relationship to ethnicity, gender, grade level, and content area. Studies were discarded from use in this meta-analysis because they failed to meet the criteria selected. With the use of meta-analysis, ethnicity, gender, grade level, and content area could be examined regardless of the age of the study, population size, participants, and geographic location of the study. Also, the four mediators listed were included as opposed to others because of their uniqueness to the students examined throughout the meta-analysis. These four factors could not be altered in any way by any of the students during the study. In addition, each of these factors was unique to each of the students. No matter if any other factors varied, these mediators remained constant. These mediators could be examined directly without needing to examine further any additional factors or mediators such as teacher efficacy, teacher certification, or years of experience.

### Mediators and Statistical Requirements

Mediators reported in this study included ethnicity, gender, grade level, and content area. Ethnicity is classified into the categories of white (non-Hispanic) and minority for the purpose of this study. Gender is classified as either male or female. Grade level refers to the year in school a

subject is in according to a traditional K-12 school setting. Content area is classified as reading, math, social studies, or science for this meta-analytic study.

In addition, statistical information, i.e. mean, median, and standard deviation, effect size for each class size reduction initiative, when available, were analyzed and used in the meta-analytical study. Furthermore, effect sizes can be categorized into one of three general types, 1) the correlation,  $r$ , 2) the standardized mean effect size,  $d$ , or 3) the odds-ratio. For the purpose of this meta-analysis, only studies that contained either the correlation,  $r$ , or the standardized mean effect size,  $d$ , were examined.

For computation and input of the correlation,  $r$ , the correlation provided in the report and the sample sizes per condition or the  $p$ -value are needed for the computer software to analyze the data. For computation of the standardized mean effect size,  $d$ , the following information must be reported for the software to analyze the data: 1) means and standard deviation per condition, per sample size, or 2) means and standard deviations and  $p$ -values, or 3) differences in means, common standard deviations and sample size, or 4) standardized mean effect size,  $d$  and sample size, or 5) mean sample size and  $t$ -values ( $t$  being the square root of  $F$  value, or 6) difference in means, sample size, and  $t$ -value, or 7) sample size and  $t$ -value only, or 8) means, sample size, and  $p$ -value, or 9) difference in means, sample size and  $p$ -value, or 10) sample size and  $p$ -value.

#### *Data Analysis*

Data were entered into a research database utilizing the *Comprehensive Meta-Analysis* (Biostat, Inc., 2000) software for the purpose of this study. Data were coded and sorted to analyze the relationship between class size reduction and student achievement in relation to the study questions. A meta-analytical approach was used in the attempt to analyze all available data pertaining to class size reduction programs.



## CHAPTER FOUR: RESEARCH OF FINDINGS

### *Overview*

This chapter presents the findings of this study. The study was designed to examine the relationship between class size and student achievement when mediated by ethnicity, gender, grade level, and content area. General class size data support the idea that smaller class size improves student achievement. Therefore, the inquiry of this study was aimed at determining what mediators had the greatest impact on student achievement. The research design was established to address the following research question:

1. What variables mediate the relationship between class size reduction and student achievement, e.g. ethnicity, gender, grade level, and content area?

A database was compiled from 26 professional educational studies. Three educational journal articles containing effect sizes pertaining to class size reduction and student achievement when mediated by ethnicity, gender, grade level, and content area were analyzed. Their mediation to class size reduction and student achievement was the basis of this study.

A review of key terms will assist in the interpretation of data results. Specific data results were examined to review the meta-analysis. Effect size is the degree to which the phenomenon is present in the population or some specific non-zero value in the population (Cooper, 1998). According to Shavelson (1996), effect size can be defined as how large a difference you want to detect in your study, not in terms of an actual score but in terms of a  $z$  score. For effect size estimation the standardized mean difference was used to estimate the effect of class size reduction on measures of student achievement. The  $d$ -index (Cohen, 1988) is a scale-free

measure of the separation between two group means. According to Cooper (1998), the *d*- index expresses the distance between the two group means in terms of their common standard deviation. Calculating the *d*-index for any comparison involves dividing the difference between the two group means by either their average standard deviation or by the standard deviation of the control group. This calculation results in a measure of the difference between the two group means expressed in terms of their common standard deviation or that of the untreated population. Thus, a *d*-index of .25 indicates that one-quarter standard deviation separates the two means. In the synthesis, the researcher subtracted the tradition class (large class) condition mean from the reduced class (small class) condition mean and divided the difference by their average standard deviation.

$$d = \frac{X_1 - X_2}{\frac{SD_1 + SD_2}{2}} ; X_1 = \text{small class mean}, X_2 = \text{large class mean}; SD_1 \text{ and } SD_2 = \text{Group Standard Deviations}$$

According to Cohen (1992), effect size can be classified as small, medium, or large. Small effect size is defined as .20 - .50. Medium effect size is defined as .50 - .80. Large effect size is defined as .80 - ∞. An effect size can be either positive or negative. A positive effect size implies a positive relationship between the analyzed factors. Conversely a negative effect size results in a negative relationship between the analyzed factors. In this study, positive effect sizes indicate that students in small classes performed better than students in traditional classes.

Another statistical value examined in this meta-analysis is the *p*-value. According to Shavelson (1996), the *p*-value describes the level of significance or is the unlikely sample result in probability terms. For example, .05 defines an unlikely event in a research study as one that

occurs five times in 100 or less. Cooper (1998) further describes  $p = .05$  as meaning only five times in 100 would sampling error create this amount of variance in effect sizes.

An additional statistical measure reported is  $Q$ . According to Cooper (1998),  $Q$  is a homogeneity statistic. The  $Q$  statistic has a chi-square distribution with  $N - 1$  degrees of freedom, or one less than the number of comparisons. The overall  $Q$ -within reported for the overall effect size indicates whether moderator tests are needed (because if variance in the samples exceeds that which the researcher would normally expect from sampling error, that is, the samples are heterogeneous, the researcher should look for moderators). The  $Q$ -between reported with the mediators indicates whether the effect sizes between the subgroups are significantly different or similar (heterogeneous or homogeneous).

Fixed and random effect sizes will also be reviewed in this meta-analysis. According to Cooper (1998), the difference between fixed and random effect sizes depends on whether the effect sizes in the data are affected by a large number of uncontrollable influences, such as differences in teachers, schools, family structures, and so on. Fixed effect sizes are based on the effect sizes obtained from the meta-analyzed studies. Random effect sizes takes into consideration hypothetical effect sizes in the efforts to generalize further the meta-analysis in the event uncontrollable influences are taken into account. For the purpose of this study, both random and fixed effect sizes were examined because there is a possibility that additional factors not analyzed in this study may have had an impact on student achievement.

The studies used in the class size reduction-student achievement meta-analysis are listed in Table 4.1. Details were included in the table including author(s) and year of publication, geographic location of the study, the number of schools, classes, students participating in each

study, grade level, content area, whether the teacher was the experimenter, and the outcome measure.

Table 1

Studies included in the Class Size Reduction-Student Achievement Meta-Analysis

Author (Year)	Type of Document	Location of Study (State)	No.Schools/ Classes/ Students	Grade Level	Content Area	Teacher as Experimenter	Outcome Measure
Finn Achilles (1999)	journal	Tennessee	nr/329/6000+	K-3	reading/ math	no/st*	effect size
				K-7	reading/ math/science social studies	no/st*	effect size
Nye Hedges Konstantopoulos (2000)	journal	Tennessee	79/nr/6000+/ males vs. females/ white vs. minority	K-3	reading/ math	no/st*	effect size
Shapson (1980)	journal	Toronto, Canada	3/62/ 16 vs. 30/ 16 vs. 37	4-5 (combined)	math concepts	no/st*	effect size

**Note: st\* = standardized test  
nr = not reported**

Table 2

Collapsed Effect Sizes Obtained from the Studies

Each effect size reflects an independent sample

FINN 1	Combined	0.254986
FINN 2	Combined	0.214988
FINN 3	Combined	0.244986
FINN 4	Combined	0.119999
FINN 5	Combined	0.184979
FINN 6	Combined	0.167472
FINN 7	Combined	0.132489
FINN K	Combined	0.164991
Nye 1M	Combined	0.309992
Nye 1W	Combined	0.199955
Nye 2M	Combined	0.239994
Nye 2W	Combined	0.129997
Nye 3M	Combined	0.239946
Nye 3W	Combined	0.124981
Nye KM	Combined	0.189981
Nye KW	Combined	0.15
Shapson	Combined	1.031471
Average Effect Size		0.179459

**Note:** 1 – 7 refers to grade level

W = white

M = minority

Stem-and-Leaf Plot Representation of Compressed Effect Sizes

STEM	LEAF
.1	1223567889
.2	13345
.3	0
.4	nr
.5	nr
.6	nr
.7	nr
.8	nr
.9	nr
1.	0

The compressed effect sizes for each study are listed in Table 4.2. More specifically, the compressed effect sizes are listed by primary author and grade level with M and W representing minority and white, respectively. Also, grades are noted by a value of K for kindergarten or a one through seven for the specific grade level. The overall average or mean effect size for all the studies is also listed in Table 4.2. The overall average or mean effect size is also noted in Figure 4.1. In reviewing the average effect sizes in Table 4.2, due to their positive numeric value, it can be stated that the mediators examined in each study had a positive effect on student achievement. The effect sizes reported ranged from .11 to 1.0. Recalling Shavelson (1996), small effect sizes equal .20 - .50, medium effect size equals .50 - .80, and large effect sizes equals .80 -  $\infty$ .

The stem-and-leaf plot serves as a graphical representation of the average effect sizes listed in Table 4.2. According to Cooper (1998), in a stem-and-leaf plot the first decimal place of each effect size acts as the stem which is placed at the left. The second decimal place acts as the leaf and is placed at the right. Leaves of effect sizes sharing the same stems are placed on the same line.

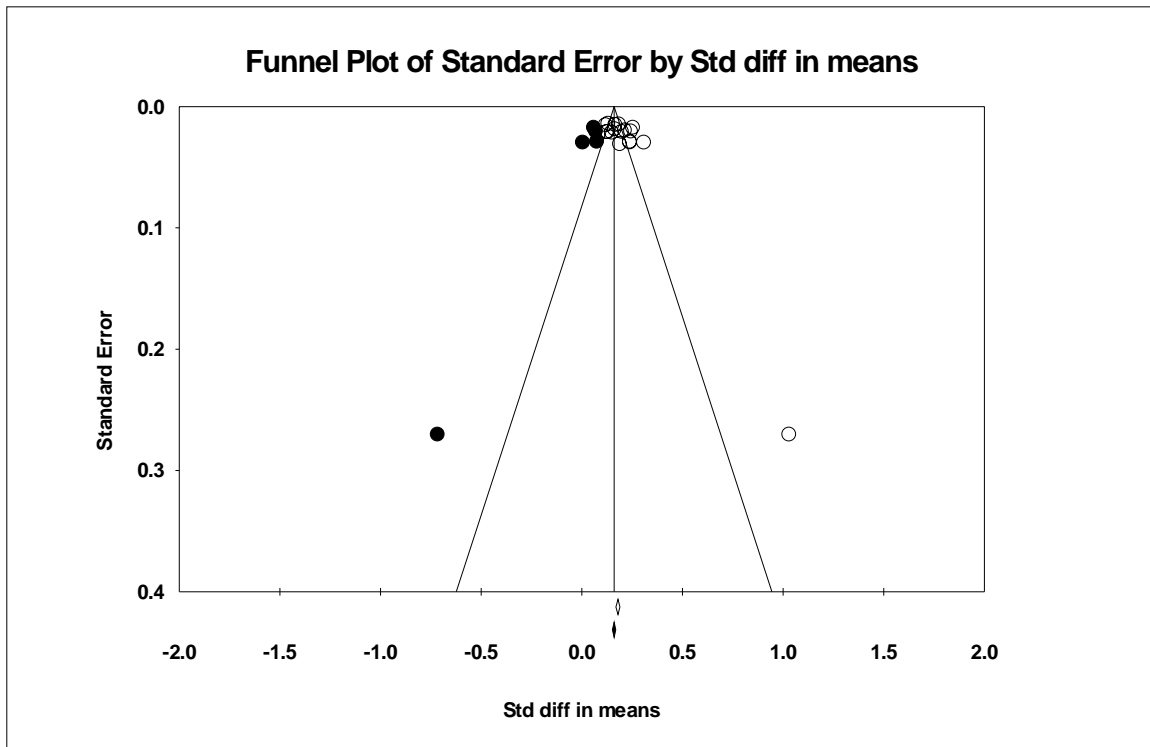


Figure 1 Funnel Plot

**Legend**

O = collapsed effect sizes collected from studies

● = theoretical effect sizes that would lead to a hypothetical normal distribution of effect sizes

*Funnel plot*

Figure 4.1 is a funnel plot. The funnel plot is a plot of a measure of study size (usually standard error or precision) on the vertical axis as a function of effect size on the horizontal axis. Studies that report a large number of effect sizes appear toward the top of the graph, and tend to cluster near the mean effect size. Studies that report a small number of effect sizes appear toward the bottom of the graph, and (since there is more sampling variation in effect size estimates in the smaller studies) will be dispersed across a range of values. This funnel plot displays the seventeen collapsed effect sizes collected from the studies used in the meta-analysis. The clear



circles displayed in Figure 4.1 represent the seventeen collapsed effect sizes collected from the studies. The darker circles represent theoretical effect sizes that would lead to a hypothetical normal distribution of effect sizes and is part of the random effects model.

The classic Fail-safe N asks if the researcher needs to be concerned that the entire observed effect may be an artifact of bias. According to Cooper (1998), the Fail-safe N answers the question, “How many findings totaling to a null hypothesis confirmation, e.g.  $z = 0$ , would have to be added to the results of the retrieved findings in order to change the conclusion that a relationship exists?” (p. 123). Rank correlation and regression procedures can test for the presence of bias. Trim and Fill offers a more nuanced perspective, and asks how the effect size would shift if the apparent bias were to be removed.

#### *Duval and Tweedie's Trim and Fill*

Table 4.3 reports the Imputed Means to the left using the fixed effects model. As noted by Cooper (1998), a fixed effects model is used when the effect sizes in a data set are not affected by a large number of uncontrollable influences such as differences in teachers, schools, family structures, and so on. If the meta-analysis had captured all the relevant studies the researcher would expect the funnel plot to be symmetric. That is, the researcher would expect studies to be dispersed equally on either side of the overall mean effect. Therefore, if the funnel plot is actually asymmetric, with a relatively high number of small studies (representing a large effect size) falling toward the right of the mean effect and relatively few falling toward the left, the researcher can be concerned that these left-hand studies may actually exist, and are missing from the analysis.

According to Borenstein and Rothstein (1999), Duval and Tweedie developed a method that allows us to impute these studies. That is, the meta-analysis software determines where the

missing studies are likely to fall, adds them to the analysis, and then recomputes the combined effect. The method is known as 'Trim and Fill' as the method initially trims the asymmetric studies from the right-hand side to locate the unbiased effect (in an iterative procedure) and then fills the plot by re-inserting the trimmed studies on the right as well as their imputed counterparts to the left the mean effect.

The Comprehensive Meta-analysis program (1998) looks for missing studies based on a fixed effect model and looks for missing studies only to the left side of the mean effect. Using these parameters, the method suggests that five studies may be missing from the meta-analysis. Under the fixed effect model the point estimate and 95% confidence interval for the combined studies, the observed values in Table 4.3, is 0.17946 (0.16996, 0.18896). Using Trim and Fill, the adjusted values in Table 4.3, the imputed point estimate is 0.16024 (0.15151, 0.16896). According to Shavelson (1996), these effect sizes can be classified as small effect sizes.

Under the random effects model the point estimate and 95% confidence interval for the combined studies, the observed values in Table 4.3, is 0.19149 (0.16513, 0.21784). Using Trim and Fill, the adjusted values in Table 4.3, the imputed point estimate is 0.16370 (0.13350, 0.19390). Similarly, according to Shavelson (1996), both of these effect sizes can be classified as small effect sizes.

Table 3: Imputed Means to the left - Fixed effects model

<b>Duval and Tweedie's trim and fill</b>								
		<b>Fixed Effects</b>			<b>Random Effects</b>			<b>Q Value</b>
	Studies Trimmed	Point Estimate	Lower Limit	Upper Limit	Point Estimate	Lower Limit	Upper Limit	
<b>Observed values</b>		0.17946	0.16996	0.18896	0.19149	0.16513	0.21784	111.71680
<b>Adjusted values</b>	5	0.16024	0.15151	0.16896	0.16370	0.13350	0.19390	223.74493

Table 4

Overall Effect of Class Size on Student Achievement

Model	Model	# of Samples	Effect size and 95% confidence interval				Test of null (2-Tail)		Homogeneity			
			Point estimate	Standard error	Variance	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value
<b>Fixed effects</b>		17	0.179	0.005	0.000	0.170	0.189	37.009	0.000	111.717	16	0.000
<b>Random effects</b>		17	0.191	0.013	0.000	0.165	0.218	14.241	0.000			

### Overall effect of class size on achievement

Table 4.4 reports the overall effect of class size on achievement by averaging 17 effect sizes, collapsed across gender, grade level, ethnicity and content area, collected from the relevant studies. This analysis indicates that for the 17 samples from three studies included in this meta-analysis, small class size has a positive effect on student achievement. The data provided here establishes an overview of the relationship between class size and student achievement as attained from the studies reviewed in the meta-analysis.

Under the fixed effect model, the weighted mean  $d$ -index for the overall effect of class size on student achievement was .179 (95 % CI = .170/.189). Under the random effects model,  $d = 0.191$  (95 % CI = .165/.218). In using Shavelson's (1996) effect size estimates, i.e. small effect = .20 - .50, medium effect = .50 - .80, and large effect = .80 -  $\infty$ , the researcher can state this analysis reports small effect sizes. More specifically, class size has a small effect on student achievement when collapsed or mediated across gender, grade level, ethnicity, and content area. Additionally, the test of homogeneity revealed  $Q(16) = 111.72, p < .0001$ .

Table 5

Mediator Results: Ethnicity

*Does ethnicity mediate the relationship between class size reduction and student achievement?*

Groups	#	Effect size and 95% confidence interval				Test of null (2-Tail)		Homogeneity		Df (Q)	P-value
		Point estimate	Standard error	Variance	Lower limit	Upper limit	Z-value	P-value	Q-value		
<b>Fixed effect analysis</b>											
Minority	4	0.246	0.015	0.000	0.217	0.275	16.452	0.000	7.945	3	0.047
White	4	0.151	0.011	0.000	0.131	0.172	14.286	0.000	7.941	3	0.047
Total within									15.885	6	0.014
Total between									26.639	1	0.000
Overall	8	0.183	0.009	0.000	0.166	0.200	21.168	0.000	42.524	7	0.000
<b>Random effect analysis</b>											
Minority	4	0.245	0.024	0.001	0.198	0.293	10.081	0.000			
White	4	0.151	0.017	0.000	0.117	0.185	8.775	0.000			
Total within											
Total between									9.949	1.000	0.002
Overall	8	0.183	0.014	0.000	0.155	0.210	12.988	0.000			

### Mediator Results: Ethnicity

A test was conducted to examine whether ethnicity mediates the relationship between class size and student achievement. Table 4.5 displays the mediated effects of ethnicity for this study. Mediator tests were conducted using both the fixed and random effects models. Four effect sizes were extracted from the studies to examine whether ethnicity, minority as opposed to white, mediates the relationship between class size and student achievement.

Under fixed-error assumptions, the weighted mean  $d$ -index was higher for minority students,  $d = .246$  (95% CI = .217/.275) than for white students,  $d = .151$  (95% CI = .131/.172),  $Q(1) = 26.64$ ,  $p < .0001$ . Using Shavelson's (1996) effect size estimates, the researcher can state minority students represented a slightly higher than small positive effect under the fixed-effects model. A small positive effect was also reported for white students under the fixed effects model.

Under the random effects model, the weighted mean  $d$ -index was also higher for minority students,  $d = .245$  (95% CI = .198/.293) compared to white students,  $d = .151$  (95% CI = .117/.185),  $Q(1) = 9.949$ ,  $p < .002$ . Under the random effects model and along with Shavelson's (1996) effect size estimates, the researcher can state minority students reported a slightly higher than small positive effect. A small positive effect was also reported for white students under the random effects model.

As indicated by the confidence intervals, using either fixed or random error models, the effect of class size on achievement was significantly, but not meaningfully, different from zero for both minority and white students. According to the data reviewed in the studies collected, it can be interpreted that ethnicity does mediate a small relationship between class size reduction and student achievement. More notably, in both the fixed and random effect models, the effect of class size reduction on achievement was greater for minority students than for white students.

Table 6

Mediator Results: Gender

*Does gender mediate the relationship between class size reduction and student achievement?*

*Small Class Advantage by Gender within Race*

SMALL VS. REGULAR CLASS COMPARISON	<u>MATH EFFECT SIZE</u>				<u>READING EFFECT SIZE</u>			
	W M	M M	W F	M F	W M	M M	W F	M F
Grade K	.23	.24	.07	.10	.18	.24	.12	.18
Grade 1	.24	.33	.22	.31	.15	.22	.19	.38
Grade 2	.12	.18	.12	.27	.15	.26	.12	.23
Grade 3	.15	.20	.06	.21	.17	.22	.12	.31

Note: Raw differences were divided by the standard deviation of the particular score across all students in all classes. W = white; M = Minority



### Mediator Results: Gender

A table was reviewed from Nye (2000) to examine whether gender mediates the relationship between class size and student achievement. Table 4.6 also displays effect sizes related to gender and ethnicity combined for grades K – 3. In addition, Table 4.6 reports data pertaining to math and reading as related to gender, ethnicity, and grades K – 3. The data reported in Table 4.7 was not included in the meta-analytic study because the data was extracted from one of the other three studies examined. This table was used due to it reporting a variety of mediators at the same time. The effect sizes are reviewed here to examine gender and its relation to class size and student achievement.

As indicated in Table 4.6, math effect sizes ranged from .06 to .33. The largest effect size, .33, was reported for Grade 1 minority males while the smallest effect size, .06, was reported for Grade 3 white females. In addition, the largest average math effect size was calculated as .23 for minority males in grades K – 3. The smallest average math effect size was calculated as .12 for white females in grades K – 3. In using Shavelson's (1996) effect size estimates, each of these effect sizes can be classified as small to slightly higher small effect sizes. The researcher can state gender mediates a small positive effect on math student achievement.

As reported in Table 4.6, reading effect sizes ranged from .12 to .38. The largest effect size, .38, was reported for Grade 1 minority females while the smallest effect size, .12, was reported for kindergarten white females, Grade 2 white females, and Grade 3 white females. In addition, the largest average reading effect size was calculated to be .27 for minority females in grades K – 3. The smallest average reading effect size was calculated as .14 for white females in grades K – 3. In using Shavelson's (1996) effect size estimates, each of these effect sizes can be

classified as small to slightly higher small effect sizes. The researcher can state gender mediates a small positive effect on reading student achievement.

According to the data reviewed in Table 4.6, it can be stated that gender does mediate the relationship between class size reduction and student achievement. The standardized test results for both male and female students at all grade levels, minority and white, reported small positive effect sizes. More notably, the effect of class size reduction on achievement was greater for minority male students in math than for white students, both male and female. Furthermore, the effect of class size reduction on achievement was greater for minority female students in reading than for white students, both male and female.

Table 7

Mediator Results: Grade level

*Does grade level mediate the relationship between class size reduction and student achievement?*

Groups Group	# Samples	Effect size and 95% confidence interval					Test of null (2-Tail)		Homogeneity		
		Point estimate	Standard error	Variance	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value
<b>Fixed effect analysis</b>											
4 thru 6	4	0.159	0.009	0.000	0.142	0.176	18.291	0.000	20.260	3	0.000
7	1	0.132	0.014	0.000	0.105	0.160	9.305	0.000	0.000	0	1.000
K-3	12	0.200	0.006	0.000	0.187	0.212	31.223	0.000	65.090	11	0.000
Total within									85.350	14	0.000
<b>Total between</b>									<b>26.367</b>	<b>2</b>	<b>0.000</b>
Overall	17	0.179	0.005	0.000	0.170	0.189	37.009	0.000	111.717	16	0.000
<b>Random effects analysis</b>											
4 thru 6	4	0.166	0.027	0.001	0.113	0.219	6.171	0.000			
7	1	0.132	0.014	0.000	0.105	0.160	9.305	0.000			
K-3	12	0.204	0.016	0.000	0.173	0.235	12.895	0.000			
Total within											
<b>Total between</b>									<b>11.242</b>	<b>2</b>	<b>0.004</b>
Overall	17	0.165	0.010	0.000	0.145	0.184	16.724	0.000			

### Mediator Results: Grade level

A test was conducted to examine whether grade level mediates the relationship between class size and student achievement. Table 4.7 displays the mediated effects of grade level for this study. Mediator tests were conducted using both the fixed and random effects models. Seventeen effect sizes were extracted from the studies to examine whether grade level, grades K – 3, 4 – 6, and 7, mediates the relationship between class size and student achievement.

Under fixed-error assumptions, the weighted mean  $d$ -index was higher for students in grades K – 3,  $d = .200$  (95% CI = .187/.212) than for students in grades 4 – 7,  $d = .159$  (95% CI = .142/.176) and students in grade 7,  $d = .132$  (95% CI = .105/.160),  $Q(2) = 26.37$ ,  $p < .001$ . In using Shavelson's (1996) effect size estimates, each of these effect sizes can be classified as small effect sizes.

Under the random effects model, the weighted mean  $d$ -index was also higher for students in grades K - 3,  $d = .204$  (95% CI = .173/.235) compared to students in grades 4 - 6,  $d = .166$  (95% CI = .113/.219) and students in grade 7,  $d = .132$  (95% CI = .105/.106),  $Q(2) = 11.242$ ,  $p < .004$ . In using Shavelson's (1996) effect size estimates, each of these effect sizes can be classified as small effect sizes.

As indicated by the confidence intervals, using either fixed or random error models, the effect of class size on achievement was significantly different, but not meaningfully so, from 0 for all grade levels analyzed. According to the data reviewed in the studies collected, it can be stated that grade level does mediate the relationship between class size reduction and student achievement. More notably, in both the fixed and random effect models, the effect of class size reduction on achievement was greater for students in grades K – 3 than for students grades 4 – 6 and 7.

Table 8

Mediator Results: Grades K-3 vs. Grades 4-6

Groups		Effect size and 95% confidence interval				Test of null (2-Tail)		Homogeneity			
Group	# Samples	Point estimate	Standard error	Variance	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value
<b>Fixed effect analysis</b>											
4 thru 6	4.00	0.159	0.009	0.000	0.142	0.176	18.291	0.000	20.260	3	0.000
K-3	12.00	0.200	0.006	0.000	0.187	0.212	31.223	0.000	65.090	11	0.000
Total within									85.350	14	0.000
Total between									14.056	1	0.000
Overall	16.00	0.186	0.005	0.000	0.176	0.196	35.991	0.000	99.405	15	0.000
<b>Random effects analysis</b>											
4 thru 6	4.00	0.166	0.027	0.001	0.113	0.219	6.171	0.000			
K-3	12.00	0.204	0.016	0.000	0.173	0.235	12.895	0.000			
Total within											
Total between									1.469	1	0.225
Overall	16.00	0.194	0.014	0.000	0.167	0.221	14.244	0.000			

### Mediator Results: K-3 vs. Grades 4-6

A test was conducted to examine whether elementary grade level, grades K – 3 and 4 – 6, mediates the relationship between class size and student achievement. Table 4.8 displays the results of this analysis. Mediator tests were conducted using both the fixed and random effects models. Sixteen effect sizes were extracted from the studies to examine whether elementary grade level, K – 3 as opposed to 4 – 6, mediates the relationship between class size and student achievement.

Under fixed-error assumptions, the weighted mean  $d$ -index was higher for students in grades K – 3,  $d = .200$  (95% CI = .187/.212) than for students in grades 4 – 6,  $d = .159$  (95% CI = .142/.176),  $Q(1) = 14.056$ ,  $p < .001$ . In using Shavelson's (1996) effect size estimates, each of these effect sizes can be classified as small effect sizes.

Under the random effects model, the weighted mean  $d$ -index was also higher for students in grades K – 3,  $d = .204$  (95% CI = .113/.219) compared to students in grades 4 – 6,  $d = .166$  (95% CI = .113/.219),  $Q(1) = 1.469$ ,  $p < .2$ . In using Shavelson's (1996) effect size estimates, each of these effect sizes can be classified as small effect sizes.

As indicated by the confidence intervals, using either fixed or random error models, the effect of class size on achievement was significantly different, but not meaningfully so, from 0 for both students in grades K – 3 and 4 – 6. According to the data reviewed in the studies collected, it can be stated that elementary grade level does mediate, in a minimal sense, the relationship between class size reduction and student achievement. More notably, in both the fixed and random effect models, the effect of class size reduction on achievement was greater for students in grades K – 3 than for students in grades 4 – 6.

Table 9

Mediator Results: K-3 vs. Grade 7

Groups Group	# Samples	Effect size and 95% confidence interval					Test of null (2-Tail) Homogeneity				
		Point estimate	Standard error	Variance	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value
<b>Fixed effect analysis</b>											
7	1	0.132	0.014	0.000	0.105	0.160	9.305	0.000	0.000	0	1.000
K-3	12	0.200	0.006	0.000	0.187	0.212	31.223	0.000	65.090	11	0.000
Total within									65.090	11	0.000
Total between									18.610	1	0.000
Overall	13	0.189	0.006	0.000	0.177	0.200	32.293	0.000	83.700	12	0.000
<b>Random effects analysis</b>											
7	1	0.132	0.014	0.000	0.105	0.160	9.305	0.000			
K-3	12	0.204	0.016	0.000	0.173	0.235	12.895	0.000			
Total within											
Total between									11.240	1	0.001
Overall	13	0.164	0.011	0.000	0.144	0.185	15.544	0.000			

### Mediator Results: Grades K-3 vs. Grade 7

A test was conducted to examine whether grade level, grades K – 3 when compared to grade 7, mediates the relationship between class size and student achievement. Table 4.9 displays the results of this analysis. Mediator tests were conducted using both the fixed and random effects models. Thirteen effect sizes were extracted from the studies to examine whether grade level, grades K – 3 as opposed to grade 7, mediates the relationship between class size and student achievement.

Under fixed-error assumptions, the weighted mean  $d$ -index was higher for students in grades K - 3,  $d = .200$  (95% CI = .187/.212) than for students in grade 7,  $d = .132$  (95% CI = .105/.160),  $Q(1) = 18.610$ ,  $p < .001$ . In using Shavelson's (1996) effect size estimates, each of these effect sizes can be classified as small effect sizes.

Under the random effects model, the weighted mean  $d$ -index was also higher for students in grades K – 3,  $d = .204$  (95% CI = .173/.235) compared to students in grade 7,  $d = .132$  (95% CI = .105/.160),  $Q(1) = 11.240$ ,  $p < .001$ . In using Shavelson's (1996) effect size estimates, each of these effect sizes can be classified as small to slightly higher small effect sizes.

As indicated by the confidence intervals, using either fixed or random error models, the effect of class size on achievement was significantly different, but not necessarily meaningfully so, from 0 for both students in grades K -3 and grade 7. According to the data reviewed in the studies collected, it can be stated that grade level, K – 3 when compared to grade 7, does mediate, in a minimal sense, the relationship between class size reduction and student achievement. More notably, in both the fixed and random effect models, the effect of class size reduction on achievement was greater for students in grades K – 3 than for students in grade 7.



Table 10

Mediator Results: Grades 4-6 vs. Grade 7

Groups Group	# Samples	Effect size and 95% confidence interval					Test of null (2-Tail) Homogeneity					
		Point estimate	Standard error	Variance	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value	
<b>Fixed effect analysis</b>												
4 thru 6	4	0.159	0.009	0.000	0.142	0.176	18.291	0.000	20.260	3	0.000	
7.00	1	0.132	0.014	0.000	0.105	0.160	9.305	0.000	0.000	0	1.000	
Total within									20.260	3	0.000	
Total between									2.582	1	0.108	
Overall	5	0.152	0.007	0.000	0.137	0.167	20.459	0.000	22.842	4	0.000	
<b>Random effects analysis</b>												
4 thru 6	4	0.166	0.027	0.001	0.113	0.219	6.171	0.000				
7.00	1	0.132	0.014	0.000	0.105	0.160	9.305	0.000				
Total within												
Total between									1.212	1	0.271	
Overall	5	0.140	0.013	0.000	0.115	0.164	11.111	0.000				

### Mediator Results: Grades 4-6 vs. Grade 7

A test was conducted to examine whether grade level, grades 4 – 6 when compared to grade 7, mediates the relationship between class size and student achievement. Table 4.10 displays the results of this analysis. Mediator tests were conducted using both the fixed and random effects models. Five effect sizes were extracted from the studies to examine whether grade level, grades 4 – 6 as opposed to grade 7, mediates the relationship between class size and student achievement.

Under fixed-error assumptions, the weighted mean  $d$ -index was higher for students in grades 4 – 6,  $d = .159$  (95% CI = .142/.176) than for students in grade 7,  $d = .132$  (95% CI = .105/.160),  $Q(1) = 2.582$ ,  $p < .2$ . In using Shavelson's (1996) effect size estimates, each of these effect sizes can be classified as small effect sizes.

Under the random effects model, the weighted mean  $d$ -index was also higher for students in grades 4 – 6,  $d = .166$  (95% CI = .113/.219) compared to students in grade 7,  $d = .132$  (95% CI = .105/.160),  $Q(1) = 1.212$ ,  $p < .3$ . In using Shavelson's (1996) effect size estimates, each of these effect sizes can be classified as small effect sizes.

As indicated by the confidence intervals, using either fixed or random error models, the effect of class size on achievement was different from 0 for both students in grades 4 – 6 and grade 7. According to the data reviewed in the studies collected, it can be stated that grade level, grades 4 – 6 when compared to grade 7, does mediate, in a minimal sense, the relationship between class size reduction and student achievement. More notably, in both the fixed and random effect models, the effect of class size reduction on achievement was greater for students in grades 4 – 6 than for students in grade 7. With the  $d$ -indexes all below .20, the mediated differences between grades 4 -6 and grade 7 are negligible.

Table 11

Mediator Results: Content Area

*Does content area mediate the relationship between class size reduction and student achievement?*

Groups Group	# Studies	Effect size and 95% confidence interval				Test of null (2-Tail)		Homogeneity			
		Point estimate	Standard error	Variance	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value
<b>Fixed effect analysis</b>											
Math	17	0.183	0.008	0.000	0.168	0.198	23.701	0.000	61.952	16	0.000
Reading	16	0.198	0.008	0.000	0.183	0.213	25.581	0.000	37.771	15	0.001
Science	4	0.145	0.015	0.000	0.116	0.175	9.783	0.000	1.441	3	0.696
social studies	4	0.132	0.015	0.000	0.103	0.161	8.894	0.000	3.797	3	0.284
Total within									104.961	37	0.000
<b>Total between</b>									<b>21.240</b>	<b>3</b>	<b>0.000</b>
Overall	41	0.179	0.005	0.000	0.170	0.189	37.009	0.000	126.201	40	0.000
<b>Random effects analysis</b>											
Math	17	0.187	0.016	0.000	0.156	0.218	11.887	0.000			
Reading	16	0.200	0.012	0.000	0.176	0.225	16.089	0.000			
Science	4	0.145	0.015	0.000	0.116	0.175	9.783	0.000			
social studies	4	0.132	0.017	0.000	0.099	0.165	7.907	0.000			
Total within											
<b>Total between</b>									<b>14.936</b>	<b>3</b>	<b>0.002</b>
Overall	41	0.171	0.007	0.000	0.157	0.185	23.312	0.000			

### Mediator Results: Content Area

A test was conducted to examine whether content area mediates the relationship between class size and student achievement. Table 4.11 displays the results of this analysis. Mediator tests were conducted using both the fixed and random effects models. Forty-one effect sizes were extracted from the studies to examine whether content area, i.e. math, reading, writing, or social studies, mediates the relationship between class size and student achievement.

Under fixed-error assumptions, the weighted mean  $d$ -index was higher for students tested in reading,  $d = .198$  (95% CI = .183/.213) than for students tested in math,  $d = .183$  (95% CI = .168/.198), students tested in science,  $d = .145$  (95% CI = .116/.175), and students tested in social studies,  $d = .132$  (95% CI = .103/.161),  $Q(3) = 21.240$ ,  $p < .001$ . In using Shavelson's (1996) effect size estimates, each of these effect sizes can be classified as small effect sizes.

Under the random effects model, the weighted mean  $d$ -index was also higher for students tested in reading,  $d = .200$  (95% CI = .176/.225) compared to students tested in math,  $d = .187$  (95% CI = .156/.218), students tested in science,  $d = .145$  (95% CI = .116/.175), and students tested in social studies,  $d = .132$  (95% CI = .099/.165),  $Q(3) = 14.936$ ,  $p < .005$ . In using Shavelson's (1996) effect size estimates, each of these effect sizes can be classified as small effect sizes.

As indicated by the confidence intervals, using either fixed or random error models, the effect of class size on achievement was different from 0 for all content areas, reading, math, science, and social studies. According to the data reviewed in the studies collected, it can be stated that content area does mediate, in a minimal sense, the relationship between class size reduction and student achievement. More notably, in both the fixed and random effect models,

the effect of class size reduction on achievement was greater for students tested in reading than for students tested in math, science, or social studies.

Table 12

Mediator Results: Math vs. Reading

Groups	Group	# Samples	Effect size and 95% confidence interval				Test of null (2-Tail)		Homogeneity			
			Point estimate	Standard error	Variance	Lower limit	Upper limit	Z-value	P-value	Q-value	Df (Q)	P-value
<b>Fixed effect analysis</b>												
	Math	17	0.183	0.008	0.000	0.168	0.198	23.701	0.000	61.952	16	0.000
	Reading	16	0.198	0.008	0.000	0.183	0.213	25.581	0.000	37.771	15	0.001
	Total within									99.723	31	0.000
	<b>Total between</b>									<b>1.801</b>	<b>1</b>	<b>0.180</b>
	Overall	33	0.190	0.005	0.000	0.180	0.201	34.847	0.000	101.524	32	0.000
<b>Random effects analysis</b>												
	Math	17	0.187	0.016	0.000	0.156	0.218	11.887	0.000			
	Reading	16	0.200	0.012	0.000	0.176	0.225	16.089	0.000			
	Total within											
	<b>Total between</b>									<b>0.425</b>	<b>1</b>	<b>0.515</b>
	Overall	33	0.195	0.010	0.000	0.176	0.214	19.993	0.000			

### Mediator Results: Math vs. Reading

A test was conducted to examine whether content area, math when compared to reading, mediates the relationship between class size and student achievement. Table 4.12 displays the results of this analysis. Mediator tests were conducted using both the fixed and random effects models. Thirty-three effect sizes were extracted from the studies to examine whether content area, math as opposed to reading, mediates the relationship between class size and student achievement.

Under fixed-error assumptions, the weighted mean  $d$ -index was higher for students tested in reading,  $d = .198$  (95% CI = .183/.213) than for students tested in math,  $d = .183$  (95% CI = .168/.198),  $Q(1) = 1.801$ ,  $p < .20$ . In using Shavelson's (1996) effect size estimates, each of these effect sizes can be classified as small effect sizes.

Under the random effects model, the weighted mean  $d$ -index was also higher for students tested in reading,  $d = .200$  (95% CI = .176/.225) compared to students tested in math,  $d = .187$  (95% CI = .156/.218),  $Q(1) = 0.425$ ,  $p < .6$ . In using Shavelson's (1996) effect size estimates, each of these effect sizes can be classified as small effect sizes.

As indicated by the confidence intervals, using either fixed or random error models, the effect of class size on achievement was different from 0, but not meaningfully so, for students tested in both reading and math. According to the data reviewed in the studies collected, it can be interpreted that content area, for both math and reading, does mediate, in a minimal sense, the relationship between class size reduction and student achievement. More notably, in both the fixed and random effect models, the effect of class size reduction on achievement was greater for students tested in reading than for students tested in math.

Table 13

Mediator Results: Math vs. Science

Groups	#	Effect size and Point estimate	95% confidence interval			Test of null (2-Tail)		Homogeneity			
			Standard error	Variance	Lower limit	Upper limit	Z-value	P-value	Q-value	Df (Q)	P-value
<b>Fixed effect analysis</b>											
Math	17	0.183	0.008	0.000	0.168	0.198	23.701	0.000	61.952	16	0.000
Science	4	0.145	0.015	0.000	0.116	0.175	9.783	0.000	1.441	3	0.696
Total within									63.393	19	0.000
<b>Total between</b>									<b>5.075</b>	<b>1</b>	<b>0.024</b>
Overall	21	0.175	0.007	0.000	0.162	0.189	25.541	0.000	68.468	20	0.000
<b>Random effects analysis</b>											
Math	17	0.187	0.016	0.000	0.156	0.218	11.887	0.000			
Science	4	0.145	0.015	0.000	0.116	0.175	9.783	0.000			
Total within											
<b>Total between</b>									<b>3.729</b>	<b>1</b>	<b>0.053</b>
Overall	21	0.165	0.011	0.000	0.144	0.186	15.273	0.000			



### Mediator Results: Math vs. Science

A test was conducted to examine whether content area, math when compared to science, mediates the relationship between class size and student achievement. Table 4.13 displays the results of this analysis. Mediator tests were conducted using both the fixed and random effects models. Twenty-one effect sizes were extracted from the studies to examine whether content area, math as opposed to science, mediates the relationship between class size and student achievement.

Under fixed-error assumptions, the weighted mean  $d$ -index was higher for students tested in math,  $d = .183$  (95% CI = .168/.198) than for students tested in science,  $d = .145$  (95% CI = .116/.175),  $Q(1) = 5.075$ ,  $p < .03$ . In using Shavelson's (1996) effect size estimates, each of these effect sizes can be classified as small effect sizes.

Under the random effects model, the weighted mean  $d$ -index was also higher for students tested in math,  $d = .187$  (95% CI = .156/.218) compared to students tested in science,  $d = .151$  (95% CI = .116/.175),  $Q(1) = 3.729$ ,  $p < .06$ . In using Shavelson's (1996) effect size estimates, each of these effect sizes can be classified as small effect sizes.

As indicated by the confidence intervals, using either fixed or random error models, the effect of class size on achievement was different from 0, but not meaningfully so, for students tested in both math and science. According to the data reviewed in the studies collected, it can be stated that content area, for both math and science, does mediate, in a minimal sense, the relationship between class size reduction and student achievement. More notably, in both the fixed and random effect models, the effect of class size reduction on achievement was greater for students tested in math than for students tested in science.

Table 14

Mediator Results: Math vs. Social Studies

Groups	#	Effect size and	95% confidence interval			Test of null (2-Tail)		Homogeneity			
Group	Samples	Point estimate	Standard error	Variance	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value
<b>Fixed effect analysis</b>											
Math	17	0.183	0.008	0.000	0.168	0.198	23.701	0.000	61.952	16	0.000
Social studies	4	0.132	0.015	0.000	0.103	0.161	8.894	0.000	3.797	3	0.284
Total within									65.749	19	0.000
<b>Total between</b>									<b>9.263</b>	<b>1</b>	<b>0.002</b>
Overall	21	0.172	0.007	0.000	0.159	0.186	25.131	0.000	75.013	20	0.000
<b>Random effects analysis</b>											
Math	17	0.187	0.016	0.000	0.156	0.218	11.887	0.000			
Social studies	4	0.132	0.017	0.000	0.099	0.165	7.907	0.000			
Total within											
<b>Total between</b>									<b>5.726</b>	<b>1</b>	<b>0.017</b>
Overall	21	0.161	0.011	0.000	0.139	0.184	14.074	0.000			

### Mediator Results: Math vs. Social Studies

A test was conducted to examine whether content area, math when compared to social studies, mediates the relationship between class size and student achievement. Table 4.14 displays the results of this analysis. Mediator tests were conducted using both the fixed and random effects models. Twenty-one effect sizes were extracted from the studies to examine whether content area, math as opposed to social studies, mediates the relationship between class size and student achievement.

Under fixed-error assumptions, the weighted mean  $d$ -index was higher for students tested in math,  $d = .183$  (95% CI = .168/.198) than for students tested in social studies,  $d = .132$  (95% CI = .103/.161),  $Q(1) = 9.263$ ,  $p < .003$ . In using Shavelson's (1996) effect size estimates, each of these effect sizes can be classified as small effect sizes.

Under the random effects model, the weighted mean  $d$ -index was also higher for students tested in math,  $d = .187$  (95% CI = .156/.218) compared to students tested in social studies,  $d = .132$  (95% CI = .099/.165),  $Q(1) = 5.726$ ,  $p < .02$ . In using Shavelson's (1996) effect size estimates, each of these effect sizes can be classified as small effect sizes.

As indicated by the confidence intervals, using either fixed or random error models, the effect of class size on achievement was different from 0, but not meaningfully so, for students tested in both math and social studies. According to the data reviewed in the studies collected, it can be stated that content area, math when compared to social studies, does mediate, in a minimal sense, the relationship between class size reduction and student achievement. More notably, in both the fixed and random effect models, the effect of class size reduction on achievement was greater for students tested in math than for students tested in social studies.

Table 15

Mediator Results: Reading vs. Science

Groups Group	# Studies	Effect size and 95% confidence interval					Test of null (2-Tail)		Homogeneity		
		Point estimate	Standard error	Variance	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value
<b>Fixed effect analysis</b>											
Reading	16	0.198	0.008	0.000	0.183	0.213	25.581	0.000	37.771	15	0.001
Science	4	0.145	0.015	0.000	0.116	0.175	9.783	0.000	1.441	3	0.696
Total within									39.212	18	0.003
<b>Total between</b>									<b>9.785</b>	<b>1</b>	<b>0.002</b>
Overall	20	0.187	0.007	0.000	0.173	0.200	27.208	0.000	48.997	19	0.000
<b>Random effects analysis</b>											
Reading	16	0.200	0.012	0.000	0.176	0.225	16.089	0.000			
Science	4	0.145	0.015	0.000	0.116	0.175	9.783	0.000			
Total within											
<b>Total between</b>									<b>8.019</b>	<b>1</b>	<b>0.005</b>
Overall	20	0.178	0.010	0.000	0.159	0.196	18.616	0.000			

### Mediator Results: Reading vs. Science

A test was conducted to examine whether content area, reading when compared to science, mediates the relationship between class size and student achievement. Table 4.15 displays the results of this analysis. Mediator tests were conducted using both the fixed and random effects models. Twenty effect sizes were extracted from the studies to examine whether content area, reading as opposed to science, moderates the relationship between class size and student achievement.

Under fixed-error assumptions, the weighted mean  $d$ -index was higher for students tested in reading,  $d = .198$  (95% CI = .183/.213) than for students tested in science,  $d = .145$  (95% CI = .116/.175),  $Q(1) = 9.785$ ,  $p < .003$ . In using Shavelson's (1996) effect size estimates, each of these effect sizes can be classified as small effect sizes.

Under the random effects model, the weighted mean  $d$ -index was also higher for students tested in reading,  $d = .200$  (95% CI = .176/.225) compared to students tested in science,  $d = .145$  (95% CI = .116/.175),  $Q(1) = 8.019$ ,  $p < .006$ . In using Shavelson's (1996) effect size estimates, each of these effect sizes can be classified as small effect sizes.

As indicated by the confidence intervals, using either fixed or random error models, the effect of class size on achievement was different from 0, but not meaningfully so, for students tested in both reading and science. According to the data reviewed in the studies collected, it can be stated that content area, reading when compared to science, does mediate, in a minimal sense, the relationship between class size reduction and student achievement. More notably, in both the fixed and random effect models, the effect of class size reduction on achievement was greater for students tested in reading than for students tested in science.

Table 16

Mediator Results: Reading vs. Social Studies

Groups Group	Number Studies	Effect size and 95% confidence interval				Test of null (2- Tail)		Homogeneity			
		Point estimate	Standard error	Variance	Lower limit	Upper limit	Z- value	P- value	Q- value	df (Q)	P-value
<b>Fixed effect analysis</b>											
Reading	16	0.198	0.008	0.000	0.183	0.213	25.581	0.000	37.771	15	0.001
social studies	4	0.132	0.015	0.000	0.103	0.161	8.894	0.000	3.797	3	0.284
Total within									41.568	18	0.001
<b>Total between</b>									<b>15.358</b>	<b>1</b>	<b>0.000</b>
Overall	20	0.184	0.007	0.000	0.170	0.197	26.798	0.000	56.926	19	0.000
<b>Random effects analysis</b>											
Reading	16	0.200	0.012	0.000	0.176	0.225	16.089	0.000			
social studies	4	0.132	0.017	0.000	0.099	0.165	7.907	0.000			
Total within											
<b>Total between</b>									<b>10.654</b>	<b>1</b>	<b>0.001</b>
Overall	20	0.176	0.010	0.000	0.156	0.196	17.627	0.000			

### Mediator Results: Reading vs. Social Studies

A test was conducted to examine whether content area, reading when compared to social studies, mediates the relationship between class size and student achievement. Table 4.16 displays the results of this analysis. Mediator tests were conducted using both the fixed and random effects models. Twenty effect sizes were extracted from the studies to examine whether content area, reading as opposed to social studies, mediates the relationship between class size and student achievement.

Under fixed-error assumptions, the weighted mean  $d$ -index was higher for students tested in reading,  $d = .198$  (95% CI = .183/.213) than for students tested in social studies,  $d = .132$  (95% CI = .103/.161),  $Q(1) = 15.358$ ,  $p < .001$ . In using Shavelson's (1996) effect size estimates, each of these effect sizes can be classified as small effect sizes.

Under the random effects model, the weighted mean  $d$ -index was also significantly higher for students tested in reading,  $d = .200$  (95% CI = .176/.225) compared to students tested in social studies,  $d = .132$  (95% CI = .099/.165),  $Q(1) = 10.654$ ,  $p < .002$ . In using Shavelson's (1996) effect size estimates, each of these effect sizes can be classified as small effect sizes.

As indicated by the confidence intervals, using either fixed or random error models, the effect of class size on achievement was, different from 0, but not meaningfully so, for students tested in both reading and social studies. According to the data reviewed in the studies collected, it can be stated that content area, reading when compared to social studies, does mediate, in a minimal sense, the relationship between class size reduction and student achievement. More notably, in both the fixed and random effect models, the effect of class size reduction on achievement was greater for students tested in reading than for students tested in social studies.

Table 17

Mediator Results: Science vs. Social Studies

Groups Group	# Samples	Effect size and 95% confidence interval					Test of null (2-Tail) Homogeneity					
		Point estimate	Standard error	Variance	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value	
<b>Fixed effect analysis</b>												
Science	4	0.145	0.015	0.000	0.116	0.175	9.783	0.000	1.441	3	0.696	
social studies	4	0.132	0.015	0.000	0.103	0.161	8.894	0.000	3.797	3	0.284	
Total within									5.238	6	0.514	
<b>Total between</b>									<b>0.397</b>	<b>1</b>	<b>0.529</b>	
Overall	8	0.139	0.011	0.000	0.118	0.159	13.206	0.000	5.635	7	0.583	
<b>Random effects analysis</b>												
Science	4	0.145	0.015	0.000	0.116	0.175	9.783	0.000				
social studies	4	0.132	0.017	0.000	0.099	0.165	7.907	0.000				
Total within												
<b>Total between</b>									<b>0.346</b>	<b>1</b>	<b>0.557</b>	
Overall	8	0.140	0.011	0.000	0.118	0.161	12.565	0.000				



### Mediator Results: Science vs. Social Studies

A test was conducted to examine whether content area, science when compared to social studies, mediates the relationship between class size and student achievement. Table 4.17 displays the results of this analysis. Mediator tests were conducted using both the fixed and random effects models. Eight effect sizes were extracted from the studies to examine whether content area, science as opposed to social studies, mediates the relationship between class size and student achievement.

Under fixed-error assumptions, the weighted mean  $d$ -index was higher for students tested in science,  $d = .145$  (95% CI = .116/.175) than for students tested in social studies,  $d = .132$  (95% CI = .103/.161),  $Q(1) = 0.397$ ,  $p < .6$ . In using Shavelson's (1996) effect size estimates, each of these effect sizes can be classified as small effect sizes.

Under the random effects model, the weighted mean  $d$ -index was also higher for students tested in science,  $d = .145$  (95% CI = .116/.175) compared to students tested in social studies,  $d = .132$  (95% CI = .099/.165),  $Q(1) = 3.46$ ,  $p < .6$ . In using Shavelson's (1996) effect size estimates, each of these effect sizes can be classified as small effect sizes.

As indicated by the confidence intervals, using either fixed or random error models, the effect of class size on achievement was different from 0, but not meaningfully so, for students tested in both science and social studies. According to the data reviewed in the studies collected, it can be stated that content area, science when compared to social studies, does mediate, in a minimal sense, the relationship between class size reduction and student achievement. More notably, in both the fixed and random effect models, the effect of class size reduction on achievement was greater for students tested in science than for students tested in social studies.

## CHAPTER FIVE: SUMMARY, DISCUSSION, RECOMMENDATIONS, AND CONCLUSION

### *Summary of Findings*

The effects of class size reduction have been studied as well as debated, both qualitatively and quantitatively, for the greater part of the 20<sup>th</sup> Century and continues now into the present day. According to Goldstein et al (2000), there is a large number of existing studies, including observational studies, matched designs and randomized controlled trials aimed at reviewing the effects of class size on student achievement.

Although class size reduction studies have been conducted, their results have varied. In fact, Flake, Von Dohlen, and Gifford (1995) claim there is no strong evidence at the national level that smaller class sizes have improved student performance. However, several studies have been conducted in the efforts to relate class size and student achievement. Glass and Smith (1979) concluded positive effects exist for classes containing less than 20 students. This evidence was based on studies classes that were considered to be “well-controlled”.

Traditional research models have been unable to combine studies with various data results. According to Goldstein et al (2000), the meta-analytical approach only requires that the reporting of studies for inclusion conforms to minimum requirements. With the selection process complete, a meta-analysis pertaining to class size and student achievement was conducted. In reviewing the data presented, effect sizes were reported as they pertained to ethnicity, gender, grade level, and content area. These areas were further examined to determine whether they mediated the relationship between class size and student achievement. Each of these constraints,

according to the meta-analysis, mediated positive effects between class size reduction and student achievement.

This study has shown, through the meta-analysis of three educational reports, that there exists a positive relationship between class size reduction and student achievement when mediated by ethnicity, gender, grade level, and content area. More specifically, minority students, when compared to white students, showed greater academic achievement in smaller classes in all content areas examined in this study. In both the fixed and random effect models, the effect of class size reduction on achievement for was greater for minority students than for white students. Also, the largest math effect size, .33, according to Shavelson (1996) can be classified as a small effect was reported for Grade 1 minority males. Similarly, the largest reading effect size, .38, according to Shavelson (1996) can be classified as a small-medium effect was reported for Grade 1 minority females.

In reviewing gender, in the one study that reported gender effect sizes, Table 4.7 reports an average effect size for males as .195, which according to Shavelson (1996) can be reported as negligible for being less than .20, while reporting an average effect size for females as .187, which can also be reported as negligible for its value less than .20. It is important to note these effect sizes are reported over Grades K – 3. In addition, these effect sizes include both white and minority students. Furthermore, these effect sizes only include effect sizes for reading and math content areas.

When reviewing the results pertaining to grade level, in both the fixed and random effect models, the effect of class size reduction on student achievement for was greater for students in grades K – 3 than for students grades 4 – 6 and 7. Although grades 4 – 6 and 7 exhibited positive effect size results, grades K – 3 proved greater with an effect size of .200, a small effect

according to Shavelson (1996). This result remained true throughout the meta-analysis even after additional review.

The content areas of math, reading, science, and social studies also revealed positive effect sizes when mediating class size and student achievement. Reading, regardless of grade level, reported the largest effect size of .198. Although larger than the other content area effect sizes, this remains negligible according to Shavelson (1996). The remaining content areas of math, science, and social studies reported effect sizes of .183, .145, and .132, respectively.

In spite of these results, few studies met the criteria for inclusion in the meta-analysis. The researcher needs to be cautious suggesting class size reduction provides positive effects on student achievement when mediated by ethnicity, gender, grade level, and content area. Instead, more research is needed on this topic and publication of that research must contain sufficient data required for meta-analysis. Considering the large amounts of money being spent on implementing class size reduction programs, not to mention the number of students, teachers, and administrators impacted by such initiatives, additional research is greatly needed.

### *Discussion*

Studies that have attempted to establish a causal link between class size reduction and student achievement have done so by using various research designs: (a) random assignment to a class with an established number of students; (b) random assignment to a class with an established number of teachers or teacher aides; (c) pre-established experiment and control groups; (d) measuring student/teacher interaction; (e) measuring classroom management and discipline time; (f) using standardized testing methods; and (g) using teacher-made assessments.

Studies that randomly assigned classrooms or students within classrooms to monitor student achievement were all flawed in some way that compromised their ability to draw strong

causal affect. Each study failed to report the class student composition, i.e. how many students were male, female, white as opposed to minority, and whether students were added to the class after the program or initiative began. This assertion leads to a generalization of results. Authors of class size studies report their findings and may not compare their results to existing studies. Therefore, additional research that reviews class size reduction and student achievement and their relationship to the mediators of gender, grade level, and content area still remains. The findings from this meta-analysis suggest a positive relationship, classified as small to small-medium, between class size and student achievement when mediated by ethnicity, gender, grade level, and content area.

The findings from the studies were consistent and encouraging, if not conclusive. These studies all indicated a positive relationship, classified as small to small-medium, between class size reduction and student achievement regardless of the mediator. Effect sizes report positive results for class size when mediated by ethnicity, gender, grade level, and content area.

#### *Limitations to Generalizability*

As mentioned, due to the research articles meeting the criteria for inclusion in this meta-analysis, it is important to note that generalizability can be difficult to achieve. Four mediators were examined in this meta-analysis. For a meta-analysis, additional mediators can also be included to examine whether additional factors play a role in mediating the relationship between class size and student achievement. It is difficult to generalize results when four mediators are examined and the possibility of others existing may also lend to the research. Also, authors to other educational journals and articles collected could have been contacted to obtain additional data that could have possibly been included in the meta-analysis.

In the studies meta-analyzed, only one form of achievement measure was reviewed, the standardized test. This may be partly due to the fact state educational policies have required student achievement be measured using standardized tests rather than teacher-made tests and student grades. It may be concluded that this measurement scale may continue largely because of the availability of standardized testing materials. In addition, by using teacher-made tests and grades, further biases may exist.

With regard to grade level, a perusal of Tables 8 – 11 suggest positive results for grades K – 7. However, only one article reported any results for grade seven. Additionally, no research pertaining to grades nine through post-secondary met the inclusion requirements for this meta-analysis. Also, each of the studies used examined only specific grade levels rather than spreading the research questions among different grade levels, say K – 12. Instead, research articles reported only results for grades K – 3.

Content area, reading and math, tended to be the more common areas examined in research articles. Only one of the articles meeting the inclusion requirements for this meta-analysis included data related to content areas other than reading and math. All of them included data results for math with two reporting data results for reading.

In addition, in recalling Figure 4.1, Funnel Plot of Standard Error by Standard Differences in Means, one would expect the studies to be distributed symmetrically about the combined effect size. By contrast, one may expect the bottom of the plot show a higher concentration of studies on one side of the mean than the other. This would reflect the fact that smaller studies (which appear toward the bottom) are more likely to be published if they have larger than average effects, which makes them more likely to meet the criterion for statistical significance. With the absence of symmetry, a generalization of the mediation that ethnicity,

gender, grade level, and content area have on class size and student achievement can be difficult to support.

Finally, not one of the studies reviewed additional academic factors that could impact standardized testing results. For example, student retentions were never mentioned. Whether students were retained during these research projects failed to be reported. This may have skewed the results for any particular grade level and content area. In addition, the academic levels of students were not addressed. Whether a student was in academically challenging classes versus remedial classes remained unmentioned.

### *Publication Bias*

It is important to note publication bias exists when conducting research. Meta-analysis is a research method aimed at combining different research results and combining them into a unified form. However, the research articles collected for this means of research may very well contain biased information and may lead to skewed results. In addition, these articles may only contain sought out information rather than a rival hypotheses or research proposal.

According to Cooper (1998), it has been known for many years that researchers whose findings conflict with the prevailing beliefs of the day are less likely to submit their results for publication. This can be based on the fact that nonsignificant findings may appear less interesting than statistically significant ones, regardless of the results provided. As an end product, research may only report one view of an educational topic rather than displaying an opposing view to contradict the hypothesis in question. Therefore, one educational view may be published more often than another.

Electronic journals, or e-journals, also contain publication bias. This method archives full-text reports using computer storage media. For example, Cooper (1998) states Internet

computer servers or compact discs have been used to store these documents for global review. However useful e-journals may be, it is important to note the concerns that arise in their collection for use. Cooper reports two characteristics that e-journals have as opposed to paper journals. First, very rarely are e-journals critiqued prior to publication. It is important for the researcher to note the differences between peer-reviewed articles and those not reviewed prior to publication to assess both the potential methodological rigor and the likelihood of publication bias. Cooper notes due to the storage capacity and favorable economics of computer technology, an article published electronically may become available to readers much more quickly than articles published in paper journals. Therefore, there is a probability that relevant research was excluded from the study at hand.

#### *Recommendations for Future Research*

Despite the positive effect sizes reported for class size and its relationship to student achievement, when mediated by ethnicity, gender, grade level, and content area, additional research is needed. Additionally, agreement on class size itself is needed in the educational arena. With the studies meta-analyzed in this study, useful information was extracted however inquiries relevant to student achievement still arise. It is important to note that additional mediators such as teacher certification, teacher efficacy, and years of experience could also be reviewed and included in such a meta-analysis. This is left to the researcher to determine in their selection criteria for inclusion.

With standardized testing gaining notoriety, a call for a national unit of measure is in line. One of the reports used for this study used a statewide standardized assessment while another state administered a different standardized test. Without knowing their assessment similarities or differences, it can be difficult to compare their results. Therefore, additional research on



measuring student achievement by means of a uniform assessment tool rather than relying on what is made available at the time of the research is needed.

Finally, an overall consensus of class size reduction and student achievement needs to be established. This can include agreement on defining class size in a uniform manner. At the very beginning of this report, class size versus class size ratio was described. However, there is little agreement in education as to which definition is more accurate and accepted. In addition, more class size research on gender and its mediation effects on class size and student achievement is needed. Only one study included data related to gender and whether it mediated class size reduction and its relationship to student achievement.

Although this study reported positive results for both male and female students, regardless of the mediator examined, with the large amounts of money being spent on class size initiatives, it is important to examine whether other initiatives can also impact student achievement at a lower cost. For example, curriculum materials, teacher training, teachers obtaining advanced education degrees, parent involvement, administrative support, and the use of differentiated instructional methods in the classroom may all also result in positive student achievement. It is difficult to conclude small class size is the only factor that will result in positive student achievement.

### *Conclusion*

Small class size and its effects on student achievement have long been debated. The findings presented in this work have contributed to the ongoing debate of whether class size reduction improves student achievement when mediated by ethnicity, gender, grade level, and content area. Small to medium-small effect sizes were reported in this meta-analysis. Although statistically significant, the results, in most cases, cannot be expressed as being meaningful to

student achievement. Due to the small effect sizes reported, the mediation of ethnicity, gender, grade level, and content area have on class size and student achievement can be considered negligible. However, it is important to note the greater effect sizes were experienced for students in grades K -3. Therefore, policymakers can incorporate the findings of this study into reviewing whether educational funding for small class size leads to a greater benefit for younger students as opposed to older students.

## REFERENCES

References marked with an asterick indicate studies included in the meta-analysis.

Armantano, Dom (2003). Let's re-think the class-size amendment [Online]. Available: [www.cato.org/dailys/08-10-03.html](http://www.cato.org/dailys/08-10-03.html)

Betts, J. and Shkolnik, J. (1999). The behavioral effects of variations in class size: The case of math teachers [Electronic version]. *Educational Evaluation and Policy Analysis*, 21, 193-213.

Borenstein, M. and Rothstein, H. (1999). Comprehensive Meta-Analysis [Computer software]. Engelwood, NJ: Biostat, Inc.

Bourke, S. (1986). How smaller is better: Some relationships between class size, teaching practices, and student achievement [Electronic version]. *American Educational Research Journal*, 23, 558-71

Brewer D., Krop C. , Gill, B., Reichardt, R. (1999). Estimating the cost of national class size reductions under different policy alternatives [Electronic version]. *Educational Evaluation and Policy Analysis*, 21, 179-92.

Chuderewicz, R (2002). Does smaller class size improve student performance? A brief review of the empirical literature [Online]. Available: [www.uidaho.edu/ucenter](http://www.uidaho.edu/ucenter)

Cohen, J. (1988). Statistical power analysis for the behavioral sciences. Hillsdale, NJ: Erlbaum.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155-159.

Cooper, H. (1998). Synthesizing research: A guide for literature reviews (3<sup>rd</sup> ed.). Thousand Oaks: SAGE Publications.

Evaluation Brief (2000). Public Schools of North Carolina, State Board of Education [Online]. Available: [www.ncpublicschools.org/Accountability/evaluation/evalbriefs/vol2n6c98.htm](http://www.ncpublicschools.org/Accountability/evaluation/evalbriefs/vol2n6c98.htm)

Fair, William (2002). Class-size amendment will cost Florida teachers dearly. *Tallahassee Democrat* [Online]. Available: [www.tallahassee.com/mld/democrat/news/opinion/4352025.htm](http://www.tallahassee.com/mld/democrat/news/opinion/4352025.htm)

Finn, J. and Achilles, C.(1990). Answers and questions about class size: A statewide experiment [Electronic version]. *American Educational Research Journal*, 27, 557-77.

- \*Finn, J. and Achilles, C. (1999). Tennessee's class size study: Findings, implications, and misconceptions [Electronic version]. *Educational Evaluation and Policy Analysis*, 21, 97-109.
- Flake, J., VonDohlen, E., Gifford, M. (1995). Class size and Student Achievement: Is there a link? [Electronic version]. *Arizona Issue Analysis*, 135, 1-27.
- Florida Department of Education (2005) [Online]. Available: [www.fldoe.org](http://www.fldoe.org)
- Florida E-Budget (2004). [Online]. Available: [www.ebudget.state.fl.us/govpriorities/education/class\\_size.asp](http://www.ebudget.state.fl.us/govpriorities/education/class_size.asp)
- Florida Tax Watch (2004). [Online]. Available: [floridataxwatch.org/archive/amendment9.html](http://floridataxwatch.org/archive/amendment9.html)
- Gilman, D. A. and Kiger, S. (2003). Should we try to keep class sizes small? [Electronic version]. *Educational Leadership*, 60, 80-85.
- Glass, G. and Smith, M. (1979). Meta-analysis of research on class size and achievement [Electronic version]. *Educational Evaluation and Policy Analysis*, 1, 2-16.
- Goldstein, H., Yang, M., Omar, R., Turner, R., and Thompson, S. (2000). Meta-analysis using multi-level models with an application to the study of class size effects [Electronic version]. *Applied Statistics*, 49, 399-412.
- Goldstein, H. and Blatchford, P. (1998). Class size and educational achievement: A review of methodology with particular reference to study designs [Electronic version]. *British Educational research Journal*, 24, 1-12.
- Good, T., Grouws, D., Mason, D., Slavings, R., & Cramer, K. (1990). An observational study of small-group mathematics instruction in elementary schools [Electronic version]. *American Educational Research Journal*, 27, 755-782.
- Haimson, L. (2000). Smaller is better: First-hand report of early grade class size reduction in New York City public schools [Electronic version]. New York, NY: Educational Priorities Panel.
- Hallinan, M. and Sorensen, B. (1985). Class size, ability group size, and student achievement [Electronic version]. *American Journal of Education*, 94, 71-89.
- Hanushek, E. A. (1998). The evidence on class size. In S.E. Mayer & P. Peterson (Eds.), *Earning and learning: How schools matter*. Washington, DC: Brookings Institution.
- JSTOR (2005). Internet-based educational research database. Available from University of Central Florida Web site, <http://www.library.ucf.edu>

- Kennedy, M. (2003). Sizing up smaller class sizes. *American School and University* [Online]. Available: [asumag.com/mag/university\\_sizing\\_smaller\\_classes](http://asumag.com/mag/university_sizing_smaller_classes)
- Krueger, A. (2000). Smaller class size: It's worth the effort [Online]. Available: [www.cgcs.org/pdfs/reducing2000.pdf](http://www.cgcs.org/pdfs/reducing2000.pdf)
- Levin, H. (1988). Cost-effectiveness and educational policy [Electronic version]. *Educational Evaluation and Policy Analysis*, 10, 51-69.
- Maxwell, N. and Lopus, J. (1995). A cost effectiveness analysis of large and small classes in the university [Electronic version]. *Educational Evaluation and Policy Analysis*, 17, 167-78.
- National Education Association (2002). Challenges facing Tennessee's public school [Online]. Available: [www.nea.org/publiced/goodnews/tn01.html](http://www.nea.org/publiced/goodnews/tn01.html)
- North Carolina Board of Education Brief (2000). [Online]. Available: [www.ncpublicschool.org/Accountability/evaluation/evalbriefs/vol2n6c98.htm](http://www.ncpublicschool.org/Accountability/evaluation/evalbriefs/vol2n6c98.htm)
- \*Nye, B.A., Hedges, L.V., & Konstantopoulos, S. (2000a). The effects of small classes on academic achievement: the results of the Tennessee class size experiment [Electronic version]. *American Educational Research Journal*, 37(1), 123-151.
- Nye, B. A., Hedges, L. V., & Konstantopoulos, S. (2000b). Do minorities and the disadvantaged benefit more from small classes? Evidence from the Tennessee class size experiment [Electronic version]. *American Journal of Education*, 109, 1-26.
- Odden, A. (1990). Class size and student achievement: Research-based policy alternatives [Electronic version]. *Educational Evaluation and Policy Analysis*, 12, 213-27.
- Pong, S. and Pallas, A. (2001). Class size and eighth-grade math achievement in the United States and abroad [Electronic version]. *Educational Evaluation and Policy Analysis*, 23, 251-73.
- PsycInfo (2005). Internet-based educational research database. Available from University of Central Florida Web site, <http://www.library.ucf.edu>
- Rice, J. (1999). The impact of class size in instructional strategies and the use of time in high school mathematics and science courses [Electronic version]. *Educational Evaluation and Policy Analysis*, 21, 215-29.
- \*Shapson, S., Wright, E., Eason, G., & Fitzgerald, J. (1980). An experimental study of the effects of class size [Electronic version]. *American Educational Research Journal*, 17, 141-152.
- Shavelson, R. (1996). Statistical reasoning for the behavioral sciences. 3<sup>rd</sup> Ed. Boston: Allyn and Bacon

- Smith, M. and Glass, G. (1980). Meta-analysis of research on class size and its relationship to attitudes and instruction [Electronic version]. *American Educational Research Journal*, 17, 419-33.
- Smith, P., Molnar, A., and Zahorik, J. (2003). Class-size reduction: A fresh look at the data [Electronic version]. *Educational leadership*, 61, 72-74.
- Southern Regional Education Board (2000). The educator supply and demand in Tennessee. Atlanta, GA: Author.
- Swennson, J. (1991). Issues in English: New NTCE policy on class size and teacher workload [Electronic version]. *The English Journal*, 80, 13-14.
- Tennessee Education Association (2002). Tennessee faces teacher shortage crisis [Online]. Available: [www.teateachers.org/thefacts/08053.php](http://www.teateachers.org/thefacts/08053.php)
- United States Department of Education (2000). The class size reduction program: Boosting student achievement in schools across the nation [Online]. Available: [www.ed.gov](http://www.ed.gov)
- Villaire, T. (n.d.). Smaller is better: Class size and student achievement [Online].
- Waymack, N. and Drury, D. (1999). Sizing it right: Class size reduction and student achievement [Electronic version]. *National School Boards Association*, 1, 1-4.
- Weglinsky (2004). The effect of class size on achievement: What the research says [Online]. Available: [www.ets.org/research/pic/memorandum.html](http://www.ets.org/research/pic/memorandum.html)